STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number 168136

TO: Jeanine Goldberg **Location: rem/2D15/2C70** **Art Unit: 1634** **Wednesday, July 13, 2005** **Case Serial Number: 10/681199**	**From: Toby Port** **Location: Biotech-Chem Library** **REM-1A59** **Phone: 571-272-2523** **toby.port@uspto.gov**

Search Notes

Examiner Goldberg,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523

STIC-Biotech/ChemLib 158136 mg

From: Goldberg, Jeanine
Sent: Tuesday, July 05, 2005 6:20 AM
To: STIC-Biotech/ChemLib
Subject: 10/681199- Search

Please search SEQ ID NO: 1 in all databases.
Please also do a mer search of SEQ ID NO: 1.

THANK YOU

Jeanine Anne Goldberg
1634
571-272-0743
REM 2D15

2C70

****
STAFF USE ONLY

Searcher:_____
Searcher Phone: 2-
Date Searcher Picked up:____
Date Completed:_____
Searcher Prep/Rev. Time:___
Online Time:_____

****
Type of Search

NA#:_____ AA#:_____
Interference:____ SPDI:___
S/L:_____ Oligomer:_____
Encode/Transl:_____
Structure#:_____ Text:___
Inventor:____ Litigation:___

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other(Specify):_____

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 04:46:25 ; Search time 758 Seconds
(without alignments)
9863.632 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1263	100.0	1263	9	ADB16964	Adb16964 Human DYX
	2	1263	100.0	1993	9	ADB16965	Adb16965 cDNA sequ
	3	1049	83.1	1641	10	ADC30210	Adc30210 Human nov
	4	1002	79.3	1263	9	ADB16939	Adb16939 Pygmy chi
	5	998	79.0	1559	11	ADM01890	Adm01890 Human cDN
	6	960	76.0	1263	9	ADB16933	Adb16933 Chimpanze
	7	859	68.0	1263	9	ADB16935	Adb16935 Gorilla D
	8	696	55.1	1263	9	ADB16937	Adb16937 Orangutan
	9	524	41.5	608	10	ADC32116	Adc32116 Human nov
	10	326	25.8	488	9	ACH35463	Ach35463 Human end
	11	270	21.4	458	9	ACH23091	Ach23091 Human adu
	12	197	15.6	1383	5	AAS70018	Aas70018 DNA encod
	13	184	14.6	49806	9	ADB16927	Adb16927 Human DYX
	14	155	12.3	49939	9	ADB16928	Adb16928 Human DYX
	15	126	10.0	164	3	AAC30498	Aac30498 Human sec
	16	109	8.6	313	3	AAC26799	Aac26799 Human sec
c	17	60	4.8	174	3	AAA45298	Aaa45298 Human sec
	18	56	4.4	174	3	AAA45298	Aaa45298 Human sec
	19	33	2.6	1697	9	ADB16924	Adb16924 cDNA sequ
	20	25	2.0	25	9	ADB16932	Adb16932 Reverse R
c	21	25	2.0	25	9	ADB16931	Adb16931 Forward R
	22	22	1.7	1316	2	AAZ07163	Aaz07163 Human lun
	23	22	1.7	1316	3	AAC79075	Aac79075 Human lun
	24	22	1.7	1316	4	AAD23150	Aad23150 Human lun
	25	22	1.7	1316	10	ADD67087	Add67087 Human lun
	26	22	1.7	1316	10	ADE87592	Ade87592 Human lun
	27	22	1.7	1431	4	ABL01953	Abl01953 Drosophil
	28	22	1.7	5746	4	ABL01952	Abl01952 Drosophil
c	29	21	1.7	579	4	AAH09057	Aah09057 Human cDN
	30	21	1.7	2310	12	ADI81619	Adi81619 C. elegan
	31	21	1.7	310268	13	ABD32548	Abd32548 Human can
	32	20	1.6	480	4	AAI86685	Aai86685 Human pol
	33	20	1.6	481	6	ABL83144	Abl83144 Human ova
c	34	20	1.6	1339	10	ADF82325	Adf82325 Leukaemia
c	35	20	1.6	2967	13	ADR08160	Adr08160 Full leng
c	36	20	1.6	3075	13	ADR08053	Adr08053 Full leng
c	37	20	1.6	3143	13	ADR66793	Adr66793 Human pro
c	38	20	1.6	3143	13	ADR65890	Adr65890 Human pro
c	39	20	1.6	3183	13	ADR08303	Adr08303 Full leng
c	40	20	1.6	8073	6	ABL32754	Abl32754 Human imm
c	41	20	1.6	9110	6	ABL34429	Abl34429 Human imm
c	42	20	1.6	48551	6	AAS20800	Aas20800 Clostridi
c	43	20	1.6	84248	6	ABQ99651	Abq99651 Human MS4
	44	20	1.6	110000	12	ADQ97138_3	Continuation (4 of
	45	20	1.6	110469	12	ADQ97337	Adq97337 Human can

ALIGNMENTS

RESULT 1
ADB16964
ID ADB16964 standard; cDNA; 1263 BP.
XX
AC ADB16964;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human DYXC1 cDNA with single nucleotide polymorphisms.
XX
KW gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW chromosome 15q21; reading disability; phonological processing;
KW rapid naming; verbal short-term memory; single nucleotide polymorphism;
KW SNP.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1263
FT /*tag= a
FT /product= "DYXC1 protein"
FT variation replace(4,t)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(572,a)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(1249,t)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(1259,g)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.

```
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16923.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 1; Page 47; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide
CC   represents the coding sequence of human DYXC1 cDNA of the invention.
XX
SQ   Sequence 1263 BP; 491 A; 214 C; 275 G; 283 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1263;  DB 9;  Length 1263;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263


RESULT 2
ADB16965
ID   ADB16965 standard; cDNA; 1993 BP.
XX
AC   ADB16965;
XX
DT   20-NOV-2003  (first entry)
XX
DE   cDNA sequence of the human DYXC1 mRNA.
XX
KW   gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW   chromosome 15q21; reading disability; phonological processing;
KW   rapid naming; verbal short-term memory; SNP;
KW   single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(205,t)
FT                   /*tag=  a
FT                   /standard_name= "Single nucleotide polymorphism"
FT   CDS             369. .1631
FT                   /*tag=  b
FT                   /product= "DYXC1 protein"
XX
```

PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.
XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
DR P-PSDB; ADB16923.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 10; Page 48-50; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC cDNA sequence of the wild type human DYXC1 mRNA of the invention.
XX
SQ Sequence 1993 BP; 693 A; 368 C; 430 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 1263; DB 9; Length 1993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631
```

RESULT 3
ADC30210
ID ADC30210 standard; cDNA; 1641 BP.
XX
AC ADC30210;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:292.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 15q21.3; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31181.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 292; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of peventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1641 BP; 614 A; 272 C; 331 G; 424 T; 0 U; 0 Other;

Query Match 83.1%; Score 1049; DB 10; Length 1641;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 163

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 223

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 283

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 343

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 403

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 463

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 523

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 583

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 643

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 703

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 763

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 823

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 883

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 943

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1003

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1063

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1123

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1124 CACAAGGCTATTGAAGATTCTTCTAAGGC 1152
```

RESULT 4
ADB16939
ID ADB16939 standard; cDNA; 1263 BP.
XX
AC ADB16939;

XX
DT 20-NOV-2003 (first entry)
XX
DE Pygmy chimpanzee DYXC1 cDNA sequence.
XX
KW gene; ss; pygmy chimpanzee; DYXC1; dyslexia; neurological disorder;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Pan paniscus.
XX
FH Key Location/Qualifiers
FT CDS 1. .1263
FT /*tag= a
FT /product= "DYXC1 protein"
XX
PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.
XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
DR P-PSDB; ADB16940.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 29; Page 126-128; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide
CC sequence is the pygmy chimpanzee DYXC1 cDNA homologous to the human DYXC1
CC gene of the invention.
XX
SQ Sequence 1263 BP; 493 A; 213 C; 274 G; 283 T; 0 U; 0 Other;

Query Match 79.3%; Score 1002; DB 9; Length 1263;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66

Qy         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126

Qy        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186

Qy        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246

Qy        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAA 306

Qy        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366

Qy        367 CGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        367 CGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426

Qy        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGG 486
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAAGCCTGG 486

Qy        487 AAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 AAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGT 546

Qy        547 CAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGA 606
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        547 CAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGA 606

Qy        607 AATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAG 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 AATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAG 666

Qy        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726

Qy        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786

Qy        787 CTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CTGCACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846

Qy        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906

Qy        907 ACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966

Qy        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026

Qy       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086

Qy       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146

Qy       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206

Qy       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263
```

RESULT 5
ADM01890
ID ADM01890 standard; cDNA; 1559 BP.

```
XX
AC   ADM01890;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human cDNA of the invention SEQ ID NO:575.
XX
KW   ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS   Homo sapiens.
XX
PN   EP1347046-A1.
XX
PD   24-SEP-2003.
XX
PF   12-APR-2002; 2002EP-00008400.
XX
PR   22-MAR-2002; 2002JP-00137785.
XX
PA   (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI   Isogai T,  Sugiyama T,  Otsuki T,  Wakamatsu A,  Sato H,  Ishii S;
PI   Yamamoto J,  Isono Y,  Hio Y,  Otsuka K,  Nagai K,  Irie R,  Tamechika I;
PI   Seki N,  Yoshikawa T,  Otsuka M,  Nagahari K,  Masuho Y;
XX
DR   WPI; 2003-723558/69.
DR   P-PSDB; ADM04333.
XX
PT   New polynucleotides and polypeptides are useful in gene therapy, for
PT   developing a diagnostic marker or medicines for regulating their
PT   expression and activity, or as a target of gene therapy.
XX
PS   Claim 1; SEQ ID NO 575; 305pp; English.
XX
CC   The invention relates to a novel human polynucleotide and the encoded
CC   polypeptide. A polynucleotide of the invention may have a use in gene
CC   therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC   as a primer for synthesizing the polynucleotide or as a probe for
CC   detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC   useful in gene therapy, for developing a diagnostic marker or medicines
CC   for regulating their expression and activity, or as a target of gene
CC   therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC   are useful as pharmaceutical agents. The present sequence represents a
CC   cDNA sequence of the invention.
XX
SQ   Sequence 1559 BP; 592 A; 252 C; 309 G; 406 T; 0 U; 0 Other;

  Query Match             79.0%;  Score 998;  DB 11;  Length 1559;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1048;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091


RESULT 6
ADB16933
ID   ADB16933 standard; cDNA; 1263 BP.
XX
AC   ADB16933;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Chimpanzee DYXC1 cDNA sequence.
XX
KW   gene; ss; chimpanzee; DYXC1; dyslexia; neurological disorder;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory.
XX
OS   Pan troglodytes.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1263
FT                   /*tag=  a
FT                   /product= "DYXC1 protein"
XX
PN   WO2003068814-A1.
```

XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.
XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
DR P-PSDB; ADB16934.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 29; Page 116-118; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide
CC sequence is the chimpanzee DYXC1 cDNA homologous to the human DYXC1 gene
CC of the invention.
XX
SQ Sequence 1263 BP; 494 A; 213 C; 272 G; 284 T; 0 U; 0 Other;

Query Match 76.0%; Score 960; DB 9; Length 1263;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1210; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy        49 GTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACG 108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        49 GTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCATG 108

Qy       109 GAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       109 GAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 168

Qy       169 ATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTAT 228
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       169 ATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTAT 228

Qy       229 AAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       229 AAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288

Qy       289 CAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAA 348
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       289 CAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAA 348

Qy       349 GCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATT 408
             |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db       349 GCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATT 408

Qy       409 GAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACT 468
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       409 GAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACT 468

Qy       469 AAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAG 528
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       469 AAAGAATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAG 528

Qy       529 AGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAA 588
             |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       529 AGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAA 588

Qy       589 TATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCA 648
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       589 TATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCA 648

Qy       649 GAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGC 708
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       649 GAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGC 708

Qy       709 AGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTA 768
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       709 AGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTA 768

Qy       769 GCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGAC 828
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       769 GCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGAC 828

Qy       829 ATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAA 888
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       829 ATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAA 888

Qy       889 GGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCC 948
             |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       889 GGAAACAAATTGTTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCC 948

Qy       949 ATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAA 1008
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       949 ATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAA 1008

Qy      1009 CTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCA 1068
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1009 CTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCA 1068

Qy      1069 CCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTC 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1069 CCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTC 1128

Qy      1129 TGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGAT 1188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1129 TGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGAT 1188

Qy      1189 CCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACA 1248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1189 CCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACA 1248

Qy      1249 GAACTAAAATCTTAA 1263
             |||||||||||||||
Db      1249 GAACTAAAATCTTAA 1263
```

RESULT 7
ADB16935
ID ADB16935 standard; cDNA; 1263 BP.
XX
AC ADB16935;
XX
DT 20-NOV-2003 (first entry)
XX
DE Gorilla DYXC1 cDNA sequence.
XX
KW gene; ss; gorilla; DYXC1; dyslexia; neurological disorder;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Gorilla gorilla.
XX
FH Key Location/Qualifiers
FT CDS 1. .1263
FT /*tag= a
FT /product= "DYXC1 protein"

```
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16936.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 29; Page 119-121; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide
CC   sequence is the gorilla DYXC1 cDNA homologous to the human DYXC1 gene of
CC   the invention.
XX
SQ   Sequence 1263 BP; 491 A; 212 C; 275 G; 285 T; 0 U; 0 Other;

  Query Match           68.0%;  Score 859;  DB 9;  Length 1263;
  Best Local Similarity  99.4%;  Pred. No. 0;
  Matches 1209;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         48 GGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCAC 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCAC 107

Qy        108 GGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCC 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCC 167

Qy        168 CATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTA 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTA 227

Qy        228 TAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 TAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGAT 287

Qy        288 GCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGA 347

Qy        348 AGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGAT 407
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        348 AGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGAT 407

Qy        408 TGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCAC 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 TGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCAC 467

Qy        468 TAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCA 527
              ||||| |||||||||||||||||||||||||||||||||||||||||||| |||||||

Db        468 TAAAGAATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAGAAATTCA 527

Qy        528 GAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAA 587
              |||||||||||| ||||||||||||||||||||||||||||||| |||||||||| ||||
Db        528 GAGAGAAGAGAAGTTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAATTAAA 587

Qy        588 ATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTC 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 ATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTC 647

Qy        648 AGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGG 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 AGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGG 707

Qy        708 CAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGT 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 CAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGT 767

Qy        768 AGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGA 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 AGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGA 827

Qy        828 CATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAA 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        828 CATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAA 887

Qy        888 AGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGC 947
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        888 AGGAAACAAATTGTTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGC 947

Qy        948 CATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAA 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        948 CATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAA 1007

Qy       1008 ACTAAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCC 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ACTAAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCC 1067

Qy       1068 ACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATT 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 ACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATT 1127

Qy       1128 CTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGA 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 CTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGA 1187

Qy       1188 TCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAAC 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 TCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAAC 1247

Qy       1248 AGAACTAAAATCTTAA 1263
              ||||||||||||||||
Db       1248 AGAACTAAAATCTTAA 1263


RESULT 8
ADB16937
ID   ADB16937 standard; cDNA; 1263 BP.
XX
AC   ADB16937;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Orangutan DYXC1 cDNA sequence.
XX
KW   gene; ss; orangutan; DYXC1; dyslexia; neurological disorder;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory.
XX
OS   Pongo pygmaeus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1263
```

```
FT                   /*tag=  a
FT                   /product= "DYXC1 protein"
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16938.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 29; Page 123-124; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide
CC   sequence is the orangutan DYXC1 cDNA homologous to the human DYXC1 gene
CC   of the invention.
XX
SQ   Sequence 1263 BP; 493 A; 218 C; 272 G; 280 T; 0 U; 0 Other;

  Query Match             55.1%;  Score 696;  DB 9;  Length 1263;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1246;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66

Qy         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126

Qy        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186

Qy        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246

Qy        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAA 306
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGACGATGCAAAGAATTAGAGAAAAA 306

Qy        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366

Qy        367 CGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        367 CGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426

Qy        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGG 486
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAAGCCTGG 486

Qy        487 AAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGT 546
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        487 AAAGAATATCAAAGAAAAGCTGAGGAGCACAAAAAAATTCAGAGAGAAGAGAAATTATGT 546

Qy        547 CAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGA 606
              |||||||||||||||||||||||||||| |||||||||| ||||||| ||||||||||||||
Db        547 CAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAATTAAAATACAAGAGTCTTACTAGA 606

Qy        607 AATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAG 666
              |||| |||||||||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        607 AATTCGGCATCTAGAAACCTTGCTCCAAAAGGAAGAAATTCAGAAAATATATTTACTGAG 666

Qy        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726

Qy        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786

Qy        787 CTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846

Qy        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906

Qy        907 ACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966

Qy        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026

Qy       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086

Qy       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146

Qy       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206

Qy       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263


RESULT 9
ADC32116
ID   ADC32116 standard; cDNA; 608 BP.
XX
AC   ADC32116;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human novel cDNA contig sequence, SEQ ID NO:2198.
XX
KW   Human; diagnostic; drug screening; forensics; gene mapping;
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
```

```
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 15q21.3; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
XX
PF   24-SEP-2002; 2002WO-US030474.
XX
PR   24-SEP-2001; 2001US-0324631P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang TY,  Zhang J,  Ren F,  Xue AJ,  Zhao QA,  Wang J,  Wehrman T;
PI   Zhou P,  Ghosh M,  Wang D,  Ma Y,  Asundi V,  Wang Z,  Weng G;
PI   Haley-Vicente D,  Drmanac RT;
XX
DR   WPI; 2003-371981/35.
DR   P-PSDB; ADC32883.
XX
PT   New polynucleotide and polypeptide useful for diagnosing, preventing or
PT   treating conditions such as neurodegenerative diseases, anemias, platelet
PT   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT   cancer.
XX
PS   Example 2; SEQ ID NO 2198; 1185pp; English.
XX
CC   The invention relates to 971 novel human cDNA sequences (ADC29919-
CC   ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC   invention also relates to nucleic acid sequences over 99% identical with
CC   the novel human cDNAs. The invention additionally encompasses expression
CC   vectors and host cells comprising a nucleic acid of the invention; the
CC   recombinant production of a polypeptide of the invention; an antibody
CC   against a polypeptide of the invention; a method of detecting
CC   polynucleotides or polypeptides of the invention; and methods of
CC   identifying a compound which binds to a polypeptide of the invention. The
CC   invention further discloses methods of peventing, treating or
CC   ameliorating a medical condition; kits comprising polynucleotide probes
CC   and/or monoclonal antibodies for carrying out the methods of the
CC   invention; methods for the identification of compounds that modulate the
CC   expression or activity of the polynucleotide and/or polypeptide; and 767
CC   contig sequences corresponding to the cDNA sequences of the invention
CC   (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC   -ADC33394). The nucleic acids and polypeptides of the invention are
CC   useful in diagnostics, drug screening, forensics, gene mapping, in the
CC   identification of mutations responsible for genetic disorders or other
CC   traits, for assessing biodiversity, and in producing many other types of
CC   data and products dependent on DNA and amino acid sequences. They are
CC   also used for treating diseases such as Parkinson's disease, Alzheimer's
CC   disease and other neurodegenerative diseases, anaemia, platelet
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a human contig
CC   sequence used in an example of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 608 BP; 230 A; 103 C; 135 G; 140 T; 0 U; 0 Other;

  Query Match             41.5%;  Score 524;  DB 10;  Length 608;
  Best Local Similarity   100.0%;  Pred. No. 1e-239;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 60

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 120

Qy        644 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 180

Qy        704 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 240

Qy        764 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 300

Qy        824 CTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGG 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGG 360

Qy        884 ATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATT 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATT 420

Qy        944 TAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACC 480

Qy       1004 TAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAG 1047
              |||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAG 524


RESULT 10
ACH35463
ID   ACH35463 standard; cDNA; 488 BP.
XX
AC   ACH35463;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human endothelial cell cDNA #3596.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 22675; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
```

```
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 488 BP; 141 A; 105 C; 126 G; 116 T; 0 U; 0 Other;

  Query Match             25.8%;  Score 326;  DB 9;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-145;
  Matches  326;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 GCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCAT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 GCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCAT 132

Qy        140 TTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTG 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTG 192

Qy        200 GGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTT 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTT 252

Qy        260 CTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAG 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 CTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAG 312

Qy        320 CACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAA 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAA 372

Qy        380 AATACGCACTAAGTGTCATGATGAAG 405
              |||||||||||||||||||||||||
Db        373 AATACGCACTAAGTGTCATGATGAAG 398


RESULT 11
ACH23091
ID   ACH23091 standard; cDNA; 458 BP.
XX
AC   ACH23091;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult ovary cDNA #1471.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 10303; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 458 BP; 170 A; 86 C; 97 G; 102 T; 0 U; 3 Other;

  Query Match             21.4%;  Score 270;  DB 9;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-118;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        773 AAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAG 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 AAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAG 248

Qy        833 CTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAA 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 CTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAA 308

Qy        893 ACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAA 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 ACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAA 368

Qy        953 GACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAA 428

Qy       1013 AAAACTTACACAAGGCTATTGAAGATTCTT 1042
              |||||||||||||||||||||||||||||
Db        429 AAAACTTACACAAGGCTATTGAAGATTCTT 458


RESULT 12
AAS70018
ID   AAS70018 standard; cDNA; 1383 BP.
XX
AC   AAS70018;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #5822.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
```

OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG05831.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 5822; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1383 BP; 399 A; 343 C; 337 G; 304 T; 0 U; 0 Other;

Query Match 15.6%; Score 197; DB 5; Length 1383;
Best Local Similarity 100.0%; Pred. No. 1.8e-83;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        172 GACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAA 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 GACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAA 768

Qy        232 AAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAA 291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 AAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAA 828

Qy        292 AGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCA 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 AGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCA 888

Qy        352 AAAGCTGCAGCAAAGCG 368
              |||||||||||||||||
Db        889 AAAGCTGCAGCAAAGCG 905
```

RESULT 13
ADB16927
ID ADB16927 standard; DNA; 49806 BP.
XX
AC ADB16927;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human DYXC1 DNA, chromosomal gene region nucleotides 50001-100000.
XX
KW ds; human; DYXC1; dyslexia; neurological disorder; chromosome 15q21;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Homo sapiens.
XX
PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.
XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 27; Page 69-83; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC partial genomic sequence of the human DYXC1 chromosomal region
CC (nucleotides 50001-100000) of the invention.
XX
SQ Sequence 49806 BP; 12975 A; 10105 C; 10577 G; 15977 T; 0 U; 172 Other;

Query Match 14.6%; Score 184; DB 9; Length 49806;
Best Local Similarity 99.6%; Pred. No. 2.8e-77;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy        404 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32444 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 32503

Qy        464 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32504 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 32563

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 583
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      32564 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAA 32623

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32624 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 32678
```

RESULT 14
ADB16928

```
ID   ADB16928 standard; DNA; 49939 BP.
XX
AC   ADB16928;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Human DYXC1 DNA, chromosomal gene region nucleotides 100001-150000.
XX
KW   ds; human; DYXC1; dyslexia; neurological disorder; chromosome 15q21;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory.
XX
OS   Homo sapiens.
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 27; Page 83-97; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC   partial genomic sequence of the human DYXC1 chromosomal region
CC   (nucleotides 100001-150000) of the invention.
XX
SQ   Sequence 49939 BP; 13786 A; 10910 C; 11127 G; 14116 T; 0 U; 0 Other;

  Query Match            12.3%;  Score 155;  DB 9;  Length 49939;
  Best Local Similarity   100.0%;  Pred. No. 2e-63;
  Matches  155;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        894 CAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAG 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14682 CAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAG 14741

Qy        954 ACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAA 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14742 ACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAA 14801

Qy       1014 AAACTTACACAAGGCTATTGAAGATTCTTCTAAGG 1048
              ||||||||||||||||||||||||||||||||||||
Db      14802 AAACTTACACAAGGCTATTGAAGATTCTTCTAAGG 14836
```

RESULT 15 AAC30498 102(b)

```
ID   AAC30498 standard; cDNA; 164 BP.
XX
AC   AAC30498;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 34573.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 34573; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 164 BP; 34 A; 43 C; 38 G; 47 T; 0 U; 2 Other;

  Query Match            10.0%;  Score 126;  DB 3;  Length 164;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-49;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 CTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGC 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGC 60

Qy         65 CCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGG 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGG 120

Qy        125 TCAACT 130
              ||||||
Db        121 TCAACT 126
```

Search completed: July 8, 2005, 07:48:42
Job time : 763 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 04:50:29 ; Search time 5660 Seconds
(without alignments)
10812.533 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1263	100.0	1993	9	AF337549	AF337549 Homo sapi
2	998	79.0	1559	6	AX833451	AX833451 Sequence
3	998	79.0	1559	9	AK095201	AK095201 Homo sapi
4	997	78.9	1468	9	BC062564	BC062564 Homo sapi
5	407	32.2	489	6	AR413597	AR413597 Sequence
6	407	32.2	489	6	AX970431	AX970431 Sequence
7	407	32.2	489	6	BD109150	BD109150 EST and e
c 8	235	18.6	114149	9	AC022083	AC022083 Homo sapi
9	214	16.9	464	6	CQ685075	CQ685075 Sequence
c 10	184	14.6	43886	2	AC021668	AC021668 Homo sapi
c 11	184	14.6	151133	9	AC013355	AC013355 Homo sapi
c 12	150	11.9	208181	2	AC016527	AC016527 Homo sapi
13	147	11.6	147	9	AY178583S2	AY178584 Pan trogl
14	147	11.6	147	9	AY178583S5	AY178587 Pan trogl
15	147	11.6	147	9	AY178592S2	AY178593 Pan panis
16	147	11.6	147	9	AY178592S5	AY178596 Pan panis
17	147	11.6	147	9	AY178601S2	AY178602 Gorilla g
18	147	11.6	147	9	AY178601S5	AY178605 Gorilla g
19	147	11.6	147	9	AY178610S2	AY178611 Pongo pyg
20	144	11.4	147	9	AY178610S5	AY178614 Pongo pyg
21	138	10.9	153	9	AY178583S7	AY178589 Pan trogl
22	138	10.9	153	9	AY178592S7	AY178598 Pan panis
23	138	10.9	153	9	AY178601S7	AY178607 Gorilla g
24	138	10.9	153	9	AY178610S7	AY178616 Pongo pyg
25	134	10.6	1187	6	CQ732459	CQ732459 Sequence
26	129	10.2	231	9	AY178583S4	AY178586 Pan trogl
27	129	10.2	231	9	AY178592S4	AY178595 Pan panis
28	126	10.0	164	6	AX918710	AX918710 Sequence
29	126	10.0	164	6	BD054243	BD054243 Sequence
30	117	9.3	123	9	AY178592S1	AY178592 Pan panis
31	117	9.3	123	9	AY178610S1	AY178610 Pongo pyg
32	113	8.9	135	9	AY178583S3	AY178585 Pan trogl
33	113	8.9	135	9	AY178592S3	AY178594 Pan panis
34	113	8.9	135	9	AY178601S3	AY178603 Gorilla g
35	111	8.8	111	9	AY178583S6	AY178588 Pan trogl
36	111	8.8	111	9	AY178583S9	AY178591 Pan trogl
37	111	8.8	111	9	AY178592S9	AY178600 Pan panis
38	111	8.8	111	9	AY178601S6	AY178606 Gorilla g
39	111	8.8	111	9	AY178601S9	AY178609 Gorilla g
40	111	8.8	111	9	AY178610S6	AY178615 Pongo pyg
41	111	8.8	111	9	AY178610S9	AY178618 Pongo pyg
42	109	8.6	313	6	AX915011	AX915011 Sequence
43	109	8.6	313	6	BD050544	BD050544 Sequence
44	105	8.3	105	9	AY178583S8	AY178590 Pan trogl
45	105	8.3	105	9	AY178592S8	AY178599 Pan panis

ALIGNMENTS

RESULT 1
AF337549
LOCUS AF337549 1993 bp mRNA linear PRI 15-JUN-2004
DEFINITION Homo sapiens EKN1 (EKN1) mRNA, complete cds.
ACCESSION AF337549
VERSION AF337549.1 GI:18478647
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1993)
AUTHORS Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B., Lyytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J., Hannula-Jouppi,K. and Kere,J.
TITLE A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
MEDLINE 22882828
PUBMED 12954984
REFERENCE 2 (bases 1 to 1993)
AUTHORS Taipale,M. and Kere,J.
TITLE A gene disrupted by translocation breakpoint in chromosome 15q21
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1993)
AUTHORS Taipale,M. and Kere,J.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2001) Finnish Genome Center, University of Helsinki, Tukholmankatu 2, Helsinki 00014, Finland
FEATURES Location/Qualifiers
source 1. .1993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21.2"
gene 1. .1993
/gene="EKN1"
/note="wildtype gene occasionally disrupted by chromosomal translocation breakpoint"
CDS 369. .1631

```
                     /gene="EKN1"
                     /codon_start=1
                     /product="EKN1"
                     /protein_id="AAL73230.1"
                     /db_xref="GI:18478648"
                     /translation="MPLQVSDYSWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKIEDMKENERIKATKALE
                     AWKEYQRKAEEQKKIQREEKLCQKEKQIKEGRKKIKYKSLTRNLASRNLAPKGRNSEN
                     IFTEKLKEDSIPAPRSVGSIKINFTPRVFPPTALRESQVAEEEEWLHKQAEARRAMNTD
                     IAELCDLKEEEKNPEWLKDKGNKLFATENYLAAINAYNLAIRLNNKMPLLYLNRAACH
                     LKLKNLHKAIEDSSKALELLMPPVTDNANARMKAHVRRGTAFCQLELYVEGLQDYEAA
                     LKIDPSNKIVQIDAEKIRNVIQGTELKS"
ORIGIN

  Query Match            100.0%;  Score 1263;  DB 9;  Length 1993;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631


RESULT 2
AX833451
LOCUS       AX833451               1559 bp    DNA     linear   PAT 15-DEC-2003
DEFINITION  Sequence 575 from Patent EP1347046.
ACCESSION   AX833451
VERSION     AX833451.1  GI:39919586
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuho,Y.
  TITLE     Full-length cDNA sequences
  JOURNAL   Patent: EP 1347046-A 575 24-SEP-2003;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            79.0%;  Score 998;  DB 6;  Length 1559;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 1048;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162
```

```
Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091
```

```
RESULT 3
AK095201
LOCUS       AK095201               1559 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ37882 fis, clone BRSTN2000536.
ACCESSION   AK095201
VERSION     AK095201.1  GI:21754405
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1559)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .1559
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRSTN2000536"
                     /tissue_type="subthalamic nucleus"
                     /clone_lib="BRSTN2"
                     /note="cloning vector: pME18SFL3"
     CDS             43. .1173
                     /note="unnamed protein product"
                     /codon_start=1
```

/protein_id="BAC04498.1"
/db_xref="GI:21754406"
/translation="MPLQVSDYSWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKIEDMKENERIKATKALE
AWKEYQRKAEEQKKIQREEKLCQKEKQIKEERKKIKYKSLTRNLASRNLAPKGRNSEN
IFTEKLKEDSIPAPRSVGSIKINFTPRVFPTALRESQVAEEEEWLHKQAEARRAMNTD
IAELCDLKEEEKNPEWLKDKGNKLFATENYLAAINAYNLAIRLNNKMPLLYLNRAACH
LKLKNLHKAIEDSSKAYRIMKRHLRLIHPTKLYKLMLRRFGM"

ORIGIN

Query Match 79.0%; Score 998; DB 9; Length 1559;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091
```

RESULT 4
BC062564
LOCUS BC062564 1468 bp mRNA linear PRI 28-NOV-2003
DEFINITION Homo sapiens dyslexia susceptibility 1 candidate 1, mRNA (cDNA clone MGC:70618 IMAGE:5163101), complete cds.
ACCESSION BC062564
VERSION BC062564.1 GI:38565951
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1468)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1468)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,

```
          Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
          McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
          Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
          Young,A., Zhang,L.-H. and Green,E.D.

          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 135 Row: a Column: 3
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 18677736.
FEATURES             Location/Qualifiers
     source          1. .1468
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:70618 IMAGE:5163101"
                     /tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             20. .1165
                     /codon_start=1
                     /product="dyslexia susceptibility 1 candidate 1"
                     /protein_id="AAH62564.1"
                     /db_xref="GI:38565952"
                     /translation="MPLQVSDYSWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKIEDMKENERIKATKALE
                     AWKEYQRKAEEQKKIQREEKLCQKEKQIKEGRKKIKYKSLTRNLASRNLAPKGRNSEN
                     IFTEKLKEDSIPAPRSVGSIKINFTPRVFPTALRESQVAEEEEWLHKQAEARRAMNTD
                     IAELCDLKEEEKNPEWLKDKGNKLFATENYLAAINAYNLAIRLNNKMPLLYLNRAVCH
                     LKLKNLHKAIEDSSKEFCSLEGIECQASEPKLSHHIPSDLHVYIQMA"
     misc_feature    38. .328
                     /note="p23; Region: homolog to co-chaperone p23"
                     /db_xref="CDD:cd00237"
ORIGIN

  Query Match             78.9%;  Score 997;  DB 9;  Length 1468;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1047;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 79

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 139

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 199

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 259

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 319

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 379

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 439

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 499

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 559

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 619

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 679

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 739

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 799

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 859

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 919

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 979

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        980 AATAAGATGCCACTATTGTATTTGAACCGGGCTGTTTGCCACCTAAAACTAAAAAACTTA 1039

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGG 1048
              ||||||||||||||||||||||||||||
Db       1040 CACAAGGCTATTGAAGATTCTTCTAAGG 1067


RESULT 5
AR413597
LOCUS       AR413597                489 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1234 from patent US 6639063.
ACCESSION   AR413597
VERSION     AR413597.1  GI:40168707
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
  TITLE     EST's and encoded human proteins
  JOURNAL   Patent: US 6639063-A 1234 28-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             32.2%;  Score 407;  DB 6;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-196;
  Matches  407;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 142

Qy         87 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 202
```

```
Qy        147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 262

Qy        207 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 322

Qy        267 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 382

Qy        327 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 442

Qy        387 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 433
              |||||||||||||||||||||||||||||||||||||||||||||
Db        443 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 489


RESULT 6
AX970431
LOCUS       AX970431                489 bp    DNA     linear   PAT 15-JAN-2004
DEFINITION  Sequence 1234 from Patent EP1104808.
ACCESSION   AX970431
VERSION     AX970431.1  GI:40977781
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
  TITLE     ESTs and encoded human proteins
  JOURNAL   Patent: EP 1104808-A 1234 06-JUN-2001;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             57. .>488
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAF15539.1"
                     /db_xref="GI:40977782"
                     /translation="MPLQVSDYXWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKI"
ORIGIN

  Query Match             32.2%;  Score 407;  DB 6;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-196;
  Matches  407;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 142

Qy         87 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 202

Qy        147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 262

Qy        207 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 322

Qy        267 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 382

Qy        327 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 442

Qy        387 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 433
              |||||||||||||||||||||||||||||||||||||||||||||
Db        443 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 489


RESULT 7
BD109150
LOCUS       BD109150                489 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  EST and encoded human protein.
ACCESSION   BD109150
VERSION     BD109150.1  GI:23203968
KEYWORDS    JP 2002010789-A/1227.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
  TITLE     EST and encoded human protein
  JOURNAL   Patent: JP 2002010789-A 1227 15-JAN-2002;
            GENSET CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002010789-A/1227
            PD   15-JAN-2002
            PF   07-AUG-2000 JP 2000280989
            PR   05-AUG-1999 US    60/147499
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE  PI
             GIORDANO
            PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
             C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
            C12N15/00
            CC   EST and encoded human protein
            FH   Key             Location/Qualifiers
            FT   CDS             57. .488.
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             32.2%;  Score 407;  DB 6;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-196;
  Matches  407;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         27 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 142

Qy         87 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 202

Qy        147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 262

Qy        207 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 322

Qy        267 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 382
```

```
Qy        327 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 442

Qy        387 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 433
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        443 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 489


RESULT 8
AC022083/c
LOCUS       AC022083               114149 bp    DNA     linear   PRI 12-MAY-2001
DEFINITION  Homo sapiens chromosome 15 clone CTD-2137J4 map 15q21.3, complete
            sequence.
ACCESSION   AC022083
VERSION     AC022083.6  GI:14029048
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 114149)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 114149)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2000) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 114149)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On May 12, 2001 this sequence version replaced gi:13621222.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399

            Note: Data from overlapping clones AC013355 [Drafting center:
            UWMSC], AC012674 [Drafting center: BCM] and AC012378 [Drafting
            center: UWMSC] were added for finishing.
FEATURES             Location/Qualifiers
     source          1. .114149
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21.3"
                     /clone="CTD-2137J4"
                     /clone_lib="Cal Tech human BAC library D"
                     /note="This clone overlaps RP11-178D12 AC013355,
                     RP11-458H3 AC012674 and RP11-420M1 AC012378. Data from
                     overlapping BACs were added and the consensus sequence
                     determined from CTD-2137J4 to the extent possible."
     misc_feature    1. .47896
                     /note="oerlap with RP11-178D12, AC013355"
     misc_feature    18530. .18640
                     /note="sequence data generated from subcloned PCR product"
     unsure          54340. .54368
                     /note="low quality data"
     unsure          90483. .90630
                     /note="low quality data"
     misc_feature    95860. .114149
                     /note="overlap with RP11-458H3, AC012674"
     misc_feature    102476. .114149
                     /note="overlap with RP11-420M1, AC012378"
ORIGIN

  Query Match            18.6%;  Score 235;  DB 9;  Length 114149;
  Best Local Similarity  100.0%;  Pred. No. 8.8e-109;
  Matches  235;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        404 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 463
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31576 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 31517

Qy        464 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 523
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31516 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 31457

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 583
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31456 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 31397

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31396 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 31342


RESULT 9
CQ685075
LOCUS       CQ685075                464 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30001 from Patent WO02070737.
ACCESSION   CQ685075
VERSION     CQ685075.1  GI:42211607
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 30001 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            16.9%;  Score 214;  DB 6;  Length 464;
  Best Local Similarity  100.0%;  Pred. No. 8.8e-98;
  Matches  214;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 108

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 168

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 228

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTG 214
```

```
                    ||||||||||||||||||||||||||||||||
Db        229 AGCAGCAAAGCAAAGATTGGGAATGACACCATTG 262


RESULT 10
AC021668/c
LOCUS       AC021668               43886 bp    DNA     linear   HTG 29-MAR-2001
DEFINITION  Homo sapiens chromosome 15 clone RP11-133O6 map 15, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC021668
VERSION     AC021668.4  GI:13488020
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43886)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-133O6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 43886)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 29, 2001 this sequence version replaced gi:11612359.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1276
                Center clone name: 133_O_6
            --------------.
            * NOTE: This record contains 53 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      745: contig of 745 bp in length
            *      746      845: gap of 100 bp
            *      846     1576: contig of 731 bp in length
            *     1577     1676: gap of 100 bp
            *     1677     2410: contig of 734 bp in length
            *     2411     2510: gap of 100 bp
            *     2511     3232: contig of 722 bp in length
            *     3233     3332: gap of 100 bp
            *     3333     4060: contig of 728 bp in length
            *     4061     4160: gap of 100 bp
            *     4161     4891: contig of 731 bp in length
            *     4892     4991: gap of 100 bp
            *     4992     5708: contig of 717 bp in length
            *     5709     5808: gap of 100 bp
            *     5809     6537: contig of 729 bp in length
            *     6538     6637: gap of 100 bp
            *     6638     7344: contig of 707 bp in length
            *     7345     7444: gap of 100 bp
            *     7445     8167: contig of 723 bp in length
            *     8168     8267: gap of 100 bp
            *     8268     8995: contig of 728 bp in length
            *     8996     9095: gap of 100 bp
            *     9096     9839: contig of 744 bp in length
            *     9840     9939: gap of 100 bp
            *     9940    10653: contig of 714 bp in length
            *    10654    10753: gap of 100 bp
            *    10754    11497: contig of 744 bp in length
            *    11498    11597: gap of 100 bp
            *    11598    12315: contig of 718 bp in length
            *    12316    12415: gap of 100 bp
            *    12416    13136: contig of 721 bp in length
            *    13137    13236: gap of 100 bp
            *    13237    13967: contig of 731 bp in length
            *    13968    14067: gap of 100 bp
            *    14068    14802: contig of 735 bp in length
            *    14803    14902: gap of 100 bp
            *    14903    15642: contig of 740 bp in length
            *    15643    15742: gap of 100 bp
            *    15743    16477: contig of 735 bp in length
            *    16478    16577: gap of 100 bp
            *    16578    17301: contig of 724 bp in length
            *    17302    17401: gap of 100 bp
            *    17402    18129: contig of 728 bp in length
            *    18130    18229: gap of 100 bp
            *    18230    18958: contig of 729 bp in length
            *    18959    19058: gap of 100 bp
            *    19059    19787: contig of 729 bp in length
            *    19788    19887: gap of 100 bp
            *    19888    20627: contig of 740 bp in length
            *    20628    20727: gap of 100 bp
            *    20728    21449: contig of 722 bp in length
            *    21450    21549: gap of 100 bp
            *    21550    22275: contig of 726 bp in length
            *    22276    22375: gap of 100 bp
            *    22376    23103: contig of 728 bp in length
            *    23104    23203: gap of 100 bp
            *    23204    23939: contig of 736 bp in length
            *    23940    24039: gap of 100 bp
            *    24040    24775: contig of 736 bp in length
            *    24776    24875: gap of 100 bp
            *    24876    25602: contig of 727 bp in length
            *    25603    25702: gap of 100 bp
            *    25703    26418: contig of 716 bp in length
            *    26419    26518: gap of 100 bp
            *    26519    27260: contig of 742 bp in length
            *    27261    27360: gap of 100 bp
            *    27361    28107: contig of 747 bp in length
            *    28108    28207: gap of 100 bp
            *    28208    28931: contig of 724 bp in length
            *    28932    29031: gap of 100 bp
            *    29032    29776: contig of 745 bp in length
            *    29777    29876: gap of 100 bp
            *    29877    30592: contig of 716 bp in length
            *    30593    30692: gap of 100 bp
            *    30693    31422: contig of 730 bp in length
            *    31423    31522: gap of 100 bp
            *    31523    32255: contig of 733 bp in length
            *    32256    32355: gap of 100 bp
            *    32356    33073: contig of 718 bp in length
            *    33074    33173: gap of 100 bp
            *    33174    33891: contig of 718 bp in length
```

```
            *    33892    33991: gap of 100 bp
            *    33992    34711: contig of 720 bp in length
            *    34712    34811: gap of 100 bp
            *    34812    35546: contig of 735 bp in length
            *    35547    35646: gap of 100 bp
            *    35647    36407: contig of 761 bp in length
            *    36408    36507: gap of 100 bp
            *    36508    37258: contig of 751 bp in length
            *    37259    37358: gap of 100 bp
            *    37359    38073: contig of 715 bp in length
            *    38074    38173: gap of 100 bp
            *    38174    38908: contig of 735 bp in length
            *    38909    39008: gap of 100 bp
            *    39009    39729: contig of 721 bp in length
            *    39730    39829: gap of 100 bp
            *    39830    40548: contig of 719 bp in length
            *    40549    40648: gap of 100 bp
            *    40649    41383: contig of 735 bp in length
            *    41384    41483: gap of 100 bp
            *    41484    42222: contig of 739 bp in length
            *    42223    42322: gap of 100 bp
            *    42323    43055: contig of 733 bp in length
            *    43056    43155: gap of 100 bp
            *    43156    43886: contig of 731 bp in length.
FEATURES             Location/Qualifiers
     source          1. .43886
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-133O6"
                     /clone_lib="RPCI-11 Human Male BAC"
ORIGIN

  Query Match             14.6%;  Score 184;  DB 2;  Length 43886;
  Best Local Similarity   99.6%;  Pred. No. 1.3e-82;
  Matches  234;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        404 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6294 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 6235

Qy        464 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6234 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 6175

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 583
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       6174 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAA 6115

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6114 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 6060


RESULT 11
AC013355/c
LOCUS       AC013355              151133 bp    DNA     linear   PRI 15-JUN-2001
DEFINITION  Homo sapiens chromosome 15 clone RP11-178D12 map 15q21.3, complete
            sequence.
ACCESSION   AC013355
VERSION     AC013355.7  GI:14456641
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 151133)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 151133)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Harrison,G., James,R., Madan,A.,
            Owen,M.P., Ratcliffe,A., Shaffer,T. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 151133)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JUN-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On Jun 15, 2001 this sequence version replaced gi:13621219.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399

            Note: Data from overlapping clones AC018926 [Drafting center:
            UWMSC] and AC022083 [Drafting center: UWMSC] were added for
            finishing.
FEATURES             Location/Qualifiers
     source          1. .151133
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21.3"
                     /clone="RP11-178D12"
                     /clone_lib="RPCI human BAC library 11"
                     /note="This clone overlaps RP11-139H15 AC018926 and
                     CTD-2137J4 AC022083. Data from overlapping BACs were added
                     and the consensus sequence determined from RP11-178D12 to
                     the extent possible."
     misc_feature    1. .53634
                     /note="overlap with RP11-139H15, AC018926"
     unsure          60513. .60518
                     /note="low quality data"
     unsure          81443. .81446
                     /note="low quality data"
     unsure          96509. .96512
                     /note="low quality data"
     misc_feature    103237. .151133
                     /note="overlap with CTD-2137J4, AC022083"
     misc_feature    121540. .121935
                     /note="sequence data generated from subcloned PCR product"
ORIGIN

  Query Match             14.6%;  Score 184;  DB 9;  Length 151133;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-82;
  Matches  234;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        404 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134810 AGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAG 134751

Qy        464 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134750 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 134691

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 583
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     134690 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAA 134631
```

```
Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     134630 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGG 134576
```

RESULT 12 102(b)
AC016527/c

```
LOCUS       AC016527               208181 bp    DNA     linear   HTG 04-JUN-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-460A24 map 15q21, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC016527
VERSION     AC016527.2  GI:8247792
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 208181)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A.,
            Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 208181)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-DEC-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT     On Jun 4, 2000 this sequence version replaced gi:6503271.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399
            ------------.
            * NOTE: This record contains 191 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      856: contig of 856 bp in length
            *      857      956: gap of unknown length
            *      957     1817: contig of 861 bp in length
            *     1818     1917: gap of unknown length
            *     1918     2890: contig of 973 bp in length
            *     2891     2990: gap of unknown length
            *     2991     3851: contig of 861 bp in length
            *     3852     3951: gap of unknown length
            *     3952     5018: contig of 1067 bp in length
            *     5019     5118: gap of unknown length
            *     5119     5991: contig of 873 bp in length
            *     5992     6091: gap of unknown length
            *     6092     7038: contig of 947 bp in length
            *     7039     7138: gap of unknown length
            *     7139     8021: contig of 883 bp in length
            *     8022     8121: gap of unknown length
            *     8122     9152: contig of 1031 bp in length
            *     9153     9252: gap of unknown length
            *     9253    10196: contig of 944 bp in length
            *    10197    10296: gap of unknown length
            *    10297    11211: contig of 915 bp in length
            *    11212    11311: gap of unknown length
            *    11312    12220: contig of 909 bp in length
            *    12221    12320: gap of unknown length
            *    12321    13554: contig of 1234 bp in length
            *    13555    13654: gap of unknown length
            *    13655    14552: contig of 898 bp in length
            *    14553    14652: gap of unknown length
            *    14653    15827: contig of 1175 bp in length
            *    15828    15927: gap of unknown length
            *    15928    16959: contig of 1032 bp in length
            *    16960    17059: gap of unknown length
            *    17060    18020: contig of 961 bp in length
            *    18021    18120: gap of unknown length
            *    18121    19061: contig of 941 bp in length
            *    19062    19161: gap of unknown length
            *    19162    20206: contig of 1045 bp in length
            *    20207    20306: gap of unknown length
            *    20307    21198: contig of 892 bp in length
            *    21199    21298: gap of unknown length
            *    21299    22311: contig of 1013 bp in length
            *    22312    22411: gap of unknown length
            *    22412    23315: contig of 904 bp in length
            *    23316    23415: gap of unknown length
            *    23416    24425: contig of 1010 bp in length
            *    24426    24525: gap of unknown length
            *    24526    25448: contig of 923 bp in length
            *    25449    25548: gap of unknown length
            *    25549    26468: contig of 920 bp in length
            *    26469    26568: gap of unknown length
            *    26569    27410: contig of 842 bp in length
            *    27411    27510: gap of unknown length
            *    27511    28462: contig of 952 bp in length
            *    28463    28562: gap of unknown length
            *    28563    29473: contig of 911 bp in length
            *    29474    29573: gap of unknown length
            *    29574    30791: contig of 1218 bp in length
            *    30792    30891: gap of unknown length
            *    30892    32425: contig of 1534 bp in length
            *    32426    32525: gap of unknown length
            *    32526    33459: contig of 934 bp in length
            *    33460    33559: gap of unknown length
            *    33560    34491: contig of 932 bp in length
            *    34492    34591: gap of unknown length
            *    34592    35495: contig of 904 bp in length
            *    35496    35595: gap of unknown length
            *    35596    36460: contig of 865 bp in length
            *    36461    36560: gap of unknown length
            *    36561    37496: contig of 936 bp in length
            *    37497    37596: gap of unknown length
            *    37597    38516: contig of 920 bp in length
            *    38517    38616: gap of unknown length
            *    38617    39598: contig of 982 bp in length
            *    39599    39698: gap of unknown length
            *    39699    40625: contig of 927 bp in length
            *    40626    40725: gap of unknown length
            *    40726    41999: contig of 1274 bp in length
            *    42000    42099: gap of unknown length
            *    42100    43057: contig of 958 bp in length
            *    43058    43157: gap of unknown length
            *    43158    44092: contig of 935 bp in length
            *    44093    44192: gap of unknown length
            *    44193    45076: contig of 884 bp in length
            *    45077    45176: gap of unknown length
            *    45177    46066: contig of 890 bp in length
            *    46067    46166: gap of unknown length
            *    46167    47021: contig of 855 bp in length
            *    47022    47121: gap of unknown length
            *    47122    48056: contig of 935 bp in length
            *    48057    48156: gap of unknown length
            *    48157    49117: contig of 961 bp in length
```

```
          *    49118    49217: gap of unknown length
          *    49218    50240: contig of 1023 bp in length
          *    50241    50340: gap of unknown length
          *    50341    51226: contig of 886 bp in length
          *    51227    51326: gap of unknown length
          *    51327    52251: contig of 925 bp in length
          *    52252    52351: gap of unknown length
          *    52352    53222: contig of 871 bp in length
          *    53223    53322: gap of unknown length
          *    53323    54237: contig of 915 bp in length
          *    54238    54337: gap of unknown length
          *    54338    55323: contig of 986 bp in length
          *    55324    55423: gap of unknown length
          *    55424    56377: contig of 954 bp in length
          *    56378    56477: gap of unknown length
          *    56478    57356: contig of 879 bp in length
          *    57357    57456: gap of unknown length
          *    57457    58363: contig of 907 bp in length
          *    58364    58463: gap of unknown length
          *    58464    59359: contig of 896 bp in length
          *    59360    59459: gap of unknown length
          *    59460    60347: contig of 888 bp in length
          *    60348    60447: gap of unknown length
          *    60448    61382: contig of 935 bp in length
          *    61383    61482: gap of unknown length
          *    61483    62469: contig of 987 bp in length
          *    62470    62569: gap of unknown length
          *    62570    63425: contig of 856 bp in length
          *    63426    63525: gap of unknown length
          *    63526    64491: contig of 966 bp in length
          *    64492    64591: gap of unknown length
          *    64592    65503: contig of 912 bp in length
          *    65504    65603: gap of unknown length
          *    65604    66496: contig of 893 bp in length
          *    66497    66596: gap of unknown length
          *    66597    67454: contig of 858 bp in length
          *    67455    67554: gap of unknown length
          *    67555    68608: contig of 1054 bp in length
          *    68609    68708: gap of unknown length
          *    68709    69566: contig of 858 bp in length
          *    69567    69666: gap of unknown length
          *    69667    70597: contig of 931 bp in length
          *    70598    70697: gap of unknown length
          *    70698    71587: contig of 890 bp in length
          *    71588    71687: gap of unknown length
          *    71688    72595: contig of 908 bp in length
          *    72596    72695: gap of unknown length
          *    72696    73669: contig of 974 bp in length
          *    73670    73769: gap of unknown length
          *    73770    74685: contig of 916 bp in length
          *    74686    74785: gap of unknown length
          *    74786    75768: contig of 983 bp in length
          *    75769    75868: gap of unknown length
          *    75869    76786: contig of 918 bp in length
          *    76787    76886: gap of unknown length
          *    76887    77743: contig of 857 bp in length
          *    77744    77843: gap of unknown length
          *    77844    79065: contig of 1222 bp in length
          *    79066    79165: gap of unknown length
          *    79166    80841: contig of 1676 bp in length
          *    80842    80941: gap of unknown length
          *    80942    81875: contig of 934 bp in length

  Query Match            11.9%;  Score 150;  DB 2;  Length 208181;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-65;
  Matches  150;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       122 AGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGA 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     22658 AGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGA 22599

Qy       182 GCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGG 241
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     22598 GCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGG 22539

Qy       242 CCATGTGGGAGACCCTTTCTGTGACGGGTG 271
            |||||||||||||||||||||||||||||
Db     22538 CCATGTGGGAGACCCTTTCTGTGACGGGTG 22509


RESULT 13
AY178583S2
LOCUS       AY178583S2              147 bp    DNA     linear   PRI 02-OCT-2003
DEFINITION  Pan troglodytes EKN1 (EKN1) gene, exon 3.
ACCESSION   AY178584
VERSION     AY178584.1  GI:27804579
KEYWORDS    .
SEGMENT     2 of 9
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 147)
  AUTHORS   Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B.,
            Lyytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J.,
            Hannula-Jouppi,K. and Kere,J.
  TITLE     A candidate gene for developmental dyslexia encodes a nuclear
            tetratricopeptide repeat domain protein dynamically regulated in
            brain
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
  MEDLINE   22882828
   PUBMED   12954984
REFERENCE   2  (bases 1 to 147)
  AUTHORS   Kaminen,N. and Kere,J.
  TITLE     EKN1 gene orthologs in chimpanzee, pygmy chimpanzee, gorilla and
            orangutan
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 147)
  AUTHORS   Kaminen,N. and Kere,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Department of Medical Genetics, University
            of Helsinki, PO Box 63, Haartmaninkatu 8, Helsinki FIN-00014,
            Finland
FEATURES             Location/Qualifiers
     source          1. .147
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     exon            1. .147
                     /gene="EKN1"
                     /number=3
ORIGIN

  Query Match            11.6%;  Score 147;  DB 9;  Length 147;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-63;
  Matches  147;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       124 GTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGC 183
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGC 60

Qy       184 AGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCC 243
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCC 120

Qy       244 ATGTGGGAGACCCTTTCTGTGACGGGT 270
            ||||||||||||||||||||||||||
Db       121 ATGTGGGAGACCCTTTCTGTGACGGGT 147


RESULT 14
AY178583S5
LOCUS       AY178583S5              147 bp    DNA     linear   PRI 02-OCT-2003
DEFINITION  Pan troglodytes EKN1 (EKN1) gene, exon 6.
ACCESSION   AY178587
VERSION     AY178587.1  GI:27804582
```

```
KEYWORDS    .
SEGMENT     5 of 9
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 147)
  AUTHORS   Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B.,
            Lyytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J.,
            Hannula-Jouppi,K. and Kere,J.
  TITLE     A candidate gene for developmental dyslexia encodes a nuclear
            tetratricopeptide repeat domain protein dynamically regulated in
            brain
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
  MEDLINE   22882828
   PUBMED   12954984
REFERENCE   2  (bases 1 to 147)
  AUTHORS   Kaminen,N. and Kere,J.
  TITLE     EKN1 gene orthologs in chimpanzee, pygmy chimpanzee, gorilla and
            orangutan
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 147)
  AUTHORS   Kaminen,N. and Kere,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Department of Medical Genetics, University
            of Helsinki, PO Box 63, Haartmaninkatu 8, Helsinki FIN-00014,
            Finland
FEATURES             Location/Qualifiers
     source          1. .147
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     exon            1. .147
                     /gene="EKN1"
                     /number=6
ORIGIN

  Query Match             11.6%;  Score 147;  DB 9;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-63;
  Matches  147;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        637 GGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCT 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCT 60

Qy        697 CGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGT 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGT 120

Qy        757 GAATCACAAGTAGCAGAAGAGGAGGAG 783
              |||||||||||||||||||||||||||
Db        121 GAATCACAAGTAGCAGAAGAGGAGGAG 147


RESULT 15
AY178592S2
LOCUS       AY178592S2              147 bp    DNA     linear   PRI 02-OCT-2003
DEFINITION  Pan paniscus EKN1 (EKN1) gene, exon 3.
ACCESSION   AY178593
VERSION     AY178593.1  GI:27804590
KEYWORDS    .
SEGMENT     2 of 9
SOURCE      Pan paniscus (pygmy chimpanzee)
  ORGANISM  Pan paniscus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 147)
  AUTHORS   Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B.,
            Lyytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J.,
            Hannula-Jouppi,K. and Kere,J.
  TITLE     A candidate gene for developmental dyslexia encodes a nuclear
            tetratricopeptide repeat domain protein dynamically regulated in
            brain
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
  MEDLINE   22882828
   PUBMED   12954984
REFERENCE   2  (bases 1 to 147)
  AUTHORS   Kaminen,N. and Kere,J.
  TITLE     EKN1 gene orthologs in chimpanzee, pygmy chimpanzee, gorilla and
            orangutan
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 147)
  AUTHORS   Kaminen,N. and Kere,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2002) Department of Medical Genetics, University
            of Helsinki, PO Box 63, Haartmaninkatu 8, Helsinki FIN-00014,
            Finland
FEATURES             Location/Qualifiers
     source          1. .147
                     /organism="Pan paniscus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9597"
     exon            1. .147
                     /gene="EKN1"
                     /number=3
ORIGIN

  Query Match             11.6%;  Score 147;  DB 9;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-63;
  Matches  147;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        124 GTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGC 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGC 60

Qy        184 AGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCC 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCC 120

Qy        244 ATGTGGGAGACCCTTTCTGTGACGGGT 270
              ||||||||||||||||||||||||||
Db        121 ATGTGGGAGACCCTTTCTGTGACGGGT 147


Search completed: July  8, 2005, 09:23:07
Job time : 5663 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 07:03:35 ; Search time 244 Seconds
(without alignments)
8469.744 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	407	32.2	489	4	US-09-621-976-1234	Sequence 1234, Ap
	2	126	10.0	164	4	US-09-513-999C-34573	Sequence 34573, A
	3	109	8.6	313	4	US-09-513-999C-30874	Sequence 30874, A
	4	22	1.7	1316	3	US-09-370-838-30	Sequence 30, Appl
	5	22	1.7	1316	4	US-09-854-133-30	Sequence 30, Appl
	6	20	1.6	273	4	US-09-248-796A-9280	Sequence 9280, Ap
	7	20	1.6	118143	4	US-09-949-016-17196	Sequence 17196, A
c	8	20	1.6	177293	4	US-09-949-016-16513	Sequence 16513, A
	9	19	1.5	216	4	US-09-248-796A-12670	Sequence 12670, A
c	10	19	1.5	359	4	US-09-513-999C-11853	Sequence 11853, A
c	11	19	1.5	425	2	US-08-343-443B-45	Sequence 45, Appl
	12	19	1.5	507	4	US-09-540-236-389	Sequence 389, App
	13	19	1.5	601	4	US-09-949-016-23803	Sequence 23803, A
	14	19	1.5	601	4	US-09-949-016-23804	Sequence 23804, A
c	15	19	1.5	601	4	US-09-949-016-51133	Sequence 51133, A
c	16	19	1.5	601	4	US-09-949-016-51134	Sequence 51134, A
	17	19	1.5	601	4	US-09-949-016-144890	Sequence 144890,
	18	19	1.5	601	4	US-09-949-016-144891	Sequence 144891,
	19	19	1.5	601	4	US-09-949-016-144892	Sequence 144892,
	20	19	1.5	601	4	US-09-949-016-170419	Sequence 170419,
	21	19	1.5	601	4	US-09-949-016-170420	Sequence 170420,
c	22	19	1.5	796	4	US-09-435-054A-19	Sequence 19, Appl
c	23	19	1.5	2371	2	US-08-343-443B-1	Sequence 1, Appli
c	24	19	1.5	2848	4	US-09-023-655-1364	Sequence 1364, Ap
c	25	19	1.5	35784	4	US-09-949-016-16785	Sequence 16785, A
c	26	19	1.5	35784	4	US-09-949-016-16786	Sequence 16786, A
	27	19	1.5	43435	4	US-09-949-016-12909	Sequence 12909, A
c	28	19	1.5	61198	4	US-09-949-016-17248	Sequence 17248, A
	29	19	1.5	62908	4	US-09-949-016-17554	Sequence 17554, A
	30	19	1.5	92407	4	US-09-596-002-36	Sequence 36, Appl
c	31	19	1.5	112114	4	US-09-949-016-17292	Sequence 17292, A
	32	19	1.5	113966	4	US-09-949-016-12277	Sequence 12277, A
	33	19	1.5	113967	4	US-09-949-016-17051	Sequence 17051, A
c	34	19	1.5	118923	4	US-09-949-016-13227	Sequence 13227, A
	35	19	1.5	146428	4	US-09-949-016-12620	Sequence 12620, A
	36	19	1.5	146438	4	US-09-949-016-12081	Sequence 12081, A
c	37	19	1.5	154605	4	US-09-949-016-11894	Sequence 11894, A
c	38	19	1.5	168104	4	US-09-949-016-12026	Sequence 12026, A
c	39	19	1.5	168105	4	US-09-949-016-16554	Sequence 16554, A
	40	19	1.5	177251	4	US-09-949-016-15841	Sequence 15841, A
c	41	19	1.5	205163	4	US-09-949-016-17009	Sequence 17009, A
	42	19	1.5	346112	4	US-09-949-016-13165	Sequence 13165, A
c	43	19	1.5	421118	4	US-09-949-016-16297	Sequence 16297, A
	44	19	1.5	640681	4	US-09-790-988-1	Sequence 1, Appli
	45	19	1.5	818128	4	US-09-949-016-14546	Sequence 14546, A

ALIGNMENTS

RESULT 1
US-09-621-976-1234
; Sequence 1234, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1234
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..488
US-09-621-976-1234

102(e)

Query Match 32.2%; Score 407; DB 4; Length 489;
Best Local Similarity 100.0%; Pred. No. 8.5e-185;
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         27 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 86
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAG 142

Qy         87 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 146
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATT 202

Qy        147 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGA 262

Qy        207 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 CACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGAC 322

Qy        267 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGA 382

Qy        327 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGC 442
```

```
Qy        387 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 433
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 ACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAG 489


RESULT 2
US-09-513-999C-34573
; Sequence 34573, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 34573
;   LENGTH: 164
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 128
;   OTHER INFORMATION: h=a or c or t
US-09-513-999C-34573

  Query Match             10.0%;  Score 126;  DB 4;  Length 164;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-50;
  Matches  126;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 CTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGC 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGC 60

Qy         65 CCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGG 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGG 120

Qy        125 TCAACT 130
              ||||||
Db        121 TCAACT 126


RESULT 3
US-09-513-999C-30874
; Sequence 30874, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 30874
;   LENGTH: 313
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 82
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 84
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 85
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 89
;   OTHER INFORMATION: v=a or c or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 92
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 250
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 283
;   OTHER INFORMATION: s=g or c
US-09-513-999C-30874

  Query Match             8.6%;  Score 109;  DB 4;  Length 313;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-42;
  Matches  109;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        940 AATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGC 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 AATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGC 159

Qy       1000 CACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGG 1048
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 CACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGG 208


RESULT 4
US-09-370-838-30
; Sequence 30, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Lodes, Michael J.
;  APPLICANT: Mohamath, Roadoh
;  APPLICANT: Secrist, Heather
;  TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
;  TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.475C1
;  CURRENT APPLICATION NUMBER: US/09/370,838
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: US 09/285,323
;  EARLIER FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 289
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 30
;   LENGTH: 1316
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-370-838-30

  Query Match             1.7%;  Score 22;  DB 3;  Length 1316;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        408 TGAAGAAGAAGAGAGGAAAAAA 429
              ||||||||||||||||||||||
Db        486 TGAAGAAGAAGAGAGGAAAAAA 507


RESULT 5
US-09-854-133-30
; Sequence 30, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: PastSEQ for Windows Version 3.0
; SEQ ID NO 30
;   LENGTH: 1316
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-854-133-30

  Query Match             1.7%;  Score 22;  DB 4;  Length 1316;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        408 TGAAGAAGAAGAGAGGAAAAAA 429
              ||||||||||||||||||||||
Db        486 TGAAGAAGAAGAGAGGAAAAAA 507


RESULT 6
US-09-248-796A-9280
; Sequence 9280, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 9280
;   LENGTH: 273
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-9280

  Query Match             1.6%;  Score 20;  DB 4;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        564 TAAAGAAGGAAGAAAAAAAA 583
              ||||||||||||||||||||
Db        113 TAAAGAAGGAAGAAAAAAAA 132


RESULT 7
US-09-949-016-17196
; Sequence 17196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17196
;   LENGTH: 118143
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17196

  Query Match             1.6%;  Score 20;  DB 4;  Length 118143;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        644 ATTCAGAAAATATATTTACT 663
              ||||||||||||||||||||
Db       9735 ATTCAGAAAATATATTTACT 9754


RESULT 8
US-09-949-016-16513/c
; Sequence 16513, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16513
;   LENGTH: 177293
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16513

  Query Match             1.6%;  Score 20;  DB 4;  Length 177293;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        548 AAAAAGAAAAGCAAATTAAA 567
              ||||||||||||||||||||
Db      44561 AAAAAGAAAAGCAAATTAAA 44542


RESULT 9
US-09-248-796A-12670
; Sequence 12670, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
```

```
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 12670
;   LENGTH: 216
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-12670

  Query Match             1.5%;  Score 19;  DB 4;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        572 GAAGAAAAAAAATAAAATA 590
              ||||||||||||||||||||
Db         63 GAAGAAAAAAAATAAAATA 81


RESULT 10
US-09-513-999C-11853/c
; Sequence 11853, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 11853
;   LENGTH: 359
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 241
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 242
;   OTHER INFORMATION: w=a or t
US-09-513-999C-11853

  Query Match             1.5%;  Score 19;  DB 4;  Length 359;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        570 AGGAAGAAAAAAAAATAAAA 588
              ||||||||||||||||||||
Db        180 AGGAAGAAAAAAAAATAAAA 162


RESULT 11
US-08-343-443B-45/c
; Sequence 45, Application US/08343443B
; Patent No. 5968734
;  GENERAL INFORMATION:
;    APPLICANT:  Aurias, Alain
;    APPLICANT:  Delattre, Olivier
;    APPLICANT:  Desmaze, Chantal
;    APPLICANT:  Melot, Thomas
;    APPLICANT:  Peter, Martine
;    APPLICANT:  Ploougastel, Beatrice
;    APPLICANT:  Thomas, Gilles
;    APPLICANT:  Zucman, Jessica
;    TITLE OF INVENTION:  NUCLEIC ACID CORRESPONDING TO A GENE OF
;    TITLE OF INVENTION:  CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
;    TITLE OF INVENTION:  TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
;    TITLE OF INVENTION:  TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
;    TITLE OF INVENTION:  TRANSLOCATIONS
;    NUMBER OF SEQUENCES:  129
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates
;      STREET:  230 South Fifteenth Street
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  AEDIT 1.0 DOS text editor
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/343,443B
;      FILING DATE:  18-NOV-1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR93/00494
;      FILING DATE:  19-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 92/06123
;      FILING DATE:  20-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiser, Gerard J.
;      REGISTRATION NUMBER:  19,763
;      REFERENCE/DOCKET NUMBER:  989.6121P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-875-8383
;      TELEFAX:  215-875-8394
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  425 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-343-443B-45

  Query Match             1.5%;  Score 19;  DB 2;  Length 425;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        565 AAAGAAGGAAGAAAAAAAA 583
              ||||||||||||||||||||
Db        275 AAAGAAGGAAGAAAAAAAA 257


RESULT 12
US-09-540-236-389
; Sequence 389, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
;  TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.2005-001
;  CURRENT APPLICATION NUMBER: US/09/540,236
;  CURRENT FILING DATE: 2000-04-04
;  NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 389
;   LENGTH: 507
;   TYPE: DNA
;   ORGANISM: M.catarrhalis
US-09-540-236-389

  Query Match             1.5%;  Score 19;  DB 4;  Length 507;
  Best Local Similarity   100.0%;  Pred. No. 54;
```

```
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        848 TAAAAGAAGAAGAAAAGAA 866
              ||||||||||||||||||
Db        383 TAAAAGAAGAAGAAAAGAA 401


RESULT 13
US-09-949-016-23803
; Sequence 23803, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23803
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-23803

  Query Match             1.5%;  Score 19;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        290 AAAGAATTAGAGAAAAATC 308
              ||||||||||||||||||
Db        180 AAAGAATTAGAGAAAAATC 198


RESULT 14
US-09-949-016-23804
; Sequence 23804, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23804
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-23804

  Query Match             1.5%;  Score 19;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        290 AAAGAATTAGAGAAAAATC 308
              ||||||||||||||||||
Db        422 AAAGAATTAGAGAAAAATC 440


RESULT 15
US-09-949-016-51133/c
; Sequence 51133, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51133
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-51133

  Query Match             1.5%;  Score 19;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        550 AAAGAAAAGCAAATTAAAG 568
              ||||||||||||||||||
Db        107 AAAGAAAAGCAAATTAAAG 89


Search completed: July  8, 2005, 12:07:45
Job time : 248 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:36:52 ; Search time 5659 Seconds
(without alignments)
10814.444 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1263	100.0	1993	9	AF337549	AF337549 Homo sapi
2	1047.4	82.9	1559	6	AX833451	AX833451 Sequence
3	1047.4	82.9	1559	9	AK095201	AK095201 Homo sapi
4	1047.2	82.9	1468	9	BC062564	BC062564 Homo sapi
5	846	67.0	1946	10	BC026462	BC026462 Mus muscu
6	485.2	38.4	2045	5	BC077575	BC077575 Xenopus l
7	432.6	34.3	489	6	AR413597	AR413597 Sequence
8	432.6	34.3	489	6	AX970431	AX970431 Sequence
9	432.6	34.3	489	6	BD109150	BD109150 EST and e
10	401.6	31.8	464	6	CQ685075	CQ685075 Sequence
11	390.2	30.9	1443	5	BC065881	BC065881 Danio rer
12	298	23.6	761	5	CR387522	CR387522 Gallus ga
13	275	21.8	1133	5	CR389904	CR389904 Gallus ga
c 14	241	19.1	114149	9	AC022083	AC022083 Homo sapi
c 15	239.4	19.0	43886	2	AC021668	AC021668 Homo sapi
c 16	239.4	19.0	151133	9	AC013355	AC013355 Homo sapi
c 17	234.6	18.6	208181	2	AC016527	AC016527 Homo sapi
18	227.8	18.0	231	9	AY178583S4	AY178586 Pan trogl
19	227.8	18.0	231	9	AY178592S4	AY178595 Pan panis
20	223	17.7	231	9	AY178601S4	AY178604 Gorilla g
21	219.8	17.4	231	9	AY178610S4	AY178613 Pongo pyg
22	178.8	14.2	313	6	AX915011	AX915011 Sequence
23	178.8	14.2	313	6	BD050544	BD050544 Sequence
24	162.8	12.9	164	6	AX918710	AX918710 Sequence
25	162.8	12.9	164	6	BD054243	BD054243 Sequence
26	151.4	12.0	153	9	AY178583S7	AY178589 Pan trogl
27	151.4	12.0	153	9	AY178592S7	AY178598 Pan panis
28	151.4	12.0	153	9	AY178601S7	AY178607 Gorilla g
29	151.4	12.0	153	9	AY178610S7	AY178616 Pongo pyg
30	147	11.6	147	9	AY178583S2	AY178584 Pan trogl
31	147	11.6	147	9	AY178583S5	AY178587 Pan trogl
32	147	11.6	147	9	AY178592S2	AY178593 Pan panis
33	147	11.6	147	9	AY178592S5	AY178596 Pan panis
34	147	11.6	147	9	AY178601S2	AY178602 Gorilla g
35	147	11.6	147	9	AY178601S5	AY178605 Gorilla g
36	147	11.6	147	9	AY178610S2	AY178611 Pongo pyg
37	145.4	11.5	147	9	AY178610S5	AY178614 Pongo pyg
38	134.4	10.6	1187	6	CQ732459	CQ732459 Sequence
39	133.4	10.6	135	9	AY178583S3	AY178585 Pan trogl
40	133.4	10.6	135	9	AY178592S3	AY178594 Pan panis
41	133.4	10.6	135	9	AY178601S3	AY178603 Gorilla g
42	131.8	10.4	135	9	AY178610S3	AY178612 Pongo pyg
43	125.4	9.9	187880	2	AC114669	AC114669 Mus muscu
44	121.4	9.6	123	9	AY178592S1	AY178592 Pan panis
45	121.4	9.6	123	9	AY178610S1	AY178610 Pongo pyg

ALIGNMENTS

RESULT 1
AF337549
LOCUS AF337549 1993 bp mRNA linear PRI 15-JUN-2004
DEFINITION Homo sapiens EKN1 (EKN1) mRNA, complete cds.
ACCESSION AF337549
VERSION AF337549.1 GI:18478647
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1993)
AUTHORS Taipale,M., Kaminen,N., Nopola-Hemmi,J., Haltia,T., Myllyluoma,B., Lyytinen,H., Muller,K., Kaaranen,M., Lindsberg,P.J., Hannula-Jouppi,K. and Kere,J.
TITLE A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11553-11558 (2003)
MEDLINE 22882828
PUBMED 12954984
REFERENCE 2 (bases 1 to 1993)
AUTHORS Taipale,M. and Kere,J.
TITLE A gene disrupted by translocation breakpoint in chromosome 15q21
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1993)
AUTHORS Taipale,M. and Kere,J.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2001) Finnish Genome Center, University of Helsinki, Tukholmankatu 2, Helsinki 00014, Finland
FEATURES Location/Qualifiers
source 1. .1993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21.2"
gene 1. .1993
/gene="EKN1"
/note="wildtype gene occasionally disrupted by chromosomal translocation breakpoint"
CDS 369. .1631

```
                     /gene="EKN1"
                     /codon_start=1
                     /product="EKN1"
                     /protein_id="AAL73230.1"
                     /db_xref="GI:18478648"
                     /translation="MPLQVSDYSWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKIEDMKENERIKATKALE
                     AWKEYQRKAEEQKKIQREEKLCQKEKQIKEGRKKIKYKSLTRNLASRNLAPKGRNSEN
                     IFTEKLKEDSIPAPRSVGSIKINFTPRVFPTALRESQVAEEEEWLHKQAEARRAMNTD
                     IAELCDLKEEEKNPEWLKDKGNKLFATENYLAAINAYNLAIRLNNKMPLLYLNRAACH
                     LKLKNLHKAIEDSSKALELLMPPVTDNANARMKAHVRRGTAFCQLELYVEGLQDYEAA
                     LKIDPSNKIVQIDAEKIRNVIQGTELKS"
ORIGIN

  Query Match            100.0%;  Score 1263;  DB 9;  Length 1993;
  Best Local Similarity  100.0%;  Pred. No. 3e-237;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631


RESULT 2
AX833451
LOCUS       AX833451               1559 bp    DNA     linear   PAT 15-DEC-2003
DEFINITION  Sequence 575 from Patent EP1347046.
ACCESSION   AX833451
VERSION     AX833451.1  GI:39919586
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuho,Y.
  TITLE     Full-length cDNA sequences
  JOURNAL   Patent: EP 1347046-A 575 24-SEP-2003;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .1559
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match            82.9%;  Score 1047.4;  DB 6;  Length 1559;
  Best Local Similarity  99.9%;  Pred. No. 5e-195;
  Matches 1048;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162
```

```
Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091
```

```
RESULT 3
AK095201
LOCUS       AK095201               1559 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ37882 fis, clone BRSTN2000536.
ACCESSION   AK095201
VERSION     AK095201.1  GI:21754405
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
            Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
            and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1559)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .1559
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="BRSTN2000536"
                     /tissue_type="subthalamic nucleus"
                     /clone_lib="BRSTN2"
                     /note="cloning vector: pME18SFL3"
     CDS             43. .1173
                     /note="unnamed protein product"
                     /codon_start=1
```

```
                     /protein_id="BAC04498.1"
                     /db_xref="GI:21754406"
                     /translation="MPLQVSDYSWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKIEDMKENERIKATKALE
                     AWKEYQRKAEEQKKIQREEKLCQKEKQIKEERKKIKYKSLTRNLASRNLAPKGRNSEN
                     IFTEKLKEDSIPAPRSVGSIKINFTPRVFPTALRESQVAEEEEWLHKQAEARRAMNTD
                     IAELCDLKEEEKNPEWLKDKGNKLFATENYLAAINAYNLAIRLNNKMPLLYLNRAACH
                     LKLKNLHKAIEDSSKAYRIMKRHLRLIHPTKLYKLMLRRFGM"
ORIGIN

  Query Match             82.9%;  Score 1047.4;  DB 9;  Length 1559;
  Best Local Similarity   99.9%;  Pred. No. 5e-195;
  Matches 1048;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091


RESULT 4
BC062564
LOCUS       BC062564                1468 bp    mRNA    linear   PRI 28-NOV-2003
DEFINITION  Homo sapiens dyslexia susceptibility 1 candidate 1, mRNA (cDNA
            clone MGC:70618 IMAGE:5163101), complete cds.
ACCESSION   BC062564
VERSION     BC062564.1  GI:38565951
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1468)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1468)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
```

```
          Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
          McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
          Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
          Young,A., Zhang,L.-H. and Green,E.D.

          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 135 Row: a Column: 3
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 18677736.
FEATURES             Location/Qualifiers
     source          1. .1468
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:70618 IMAGE:5163101"
                     /tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             20. .1165
                     /codon_start=1
                     /product="dyslexia susceptibility 1 candidate 1"
                     /protein_id="AAH62564.1"
                     /db_xref="GI:38565952"
                     /translation="MPLQVSDYSWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKIEDMKENERIKATKALE
                     AWKEYQRKAEEQKKIQREEKLCQKEKQIKEGRKKIKYKSLTRNLASRNLAPKGRNSEN
                     IFTEKLKEDSIPAPRSVGSIKINFTPRVFPTALRESQVAEEEEWLHKQAEARRAMNTD
                     IAELCDLKEEEKNPEWLKDKGNKLFATENYLAAINAYNLAIRLNNKMPLLYLNRAVCH
                     LKLKNLHKAIEDSSKEFCSLEGIECQASEPKLSHHIPSDLHVYIQMA"
     misc_feature    38. .328
                     /note="p23; Region: homolog to co-chaperone p23"
                     /db_xref="CDD:cd00237"
ORIGIN

  Query Match             82.9%;  Score 1047.2;  DB 9;  Length 1468;
  Best Local Similarity   99.7%;  Pred. No. 5.5e-195;
  Matches 1049;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 79

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 139

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 199

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 259

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 319

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 379

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 439

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 499

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        500 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 559

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        560 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 619

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        620 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 679

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        680 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 739

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        740 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 799

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        800 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 859

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        860 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 919

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        920 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 979

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        980 AATAAGATGCCACTATTGTATTTGAACCGGGCTGTTTGCCACCTAAAACTAAAAAACTTA 1039

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACT 1052
              |||||||||||||||||||||||||||| | |
Db       1040 CACAAGGCTATTGAAGATTCTTCTAAGGAATT 1071

RESULT 5
BC026462
LOCUS       BC026462                1946 bp    mRNA    linear   ROD 29-JUN-2004
DEFINITION  Mus musculus dyslexia susceptibility 1 candidate 1 homolog (human),
            mRNA (cDNA clone MGC:31465 IMAGE:4483015), complete cds.
ACCESSION   BC026462
VERSION     BC026462.1  GI:20071433
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1946)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
```

```
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1946)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 44 Row: b Column: 9
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis.
FEATURES             Location/Qualifiers
     source          1. .1946
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:31465 IMAGE:4483015"
                     /tissue_type="Mammary tumor. Metallothionien-TGF alpha
                     model. 10 month old virgin mouse. Taken by biopsy."
                     /clone_lib="NCI_CGAP_Mam1"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1946
                     /gene="Dyx1c1"
                     /note="synonym: EKN1"
                     /db_xref="LocusID:67685"
                     /db_xref="MGI:1914935"
     CDS             48. .1310
                     /gene="Dyx1c1"
                     /codon_start=1
                     /product="dyslexia susceptibility 1 candidate 1 homolog"
                     /protein_id="AAH26462.1"
                     /db_xref="GI:20071434"
                     /db_xref="LocusID:67685"
                     /db_xref="MGI:1914935"
                     /translation="MPVRVSEFSWQQTPATIFLSLPLRGVCVRDADVFCGESYLKVNF
                     PPFLFELFLYAPIDDGKSKAKIGNDTILFTLYKKEPVLWDSLSMPGVDKEMMQRIREK
                     SILQAQEKAKEATEAKAVAKREDQRYALGEMMKIEEEERKKIEDMKENERKKATSELE
                     AWKECQKKADGQKRVQRKEKPPERKQAEETKALKPRGLPRKAPPTRLPTRGRNWENIF
                     PEKLKEDRVPAPRSAGSIQISFTPRVFPTALRESQVAEEEEWLHKQAEARRAMSTDLP
                     EFFDLKEEERNPDWLKDKGNKLFATENYLAAVDAYNLAIRLNCKIPLLYLNRAACHLK
                     LKNLHKAIEDSSKALELLTPPVADNANARMKAHVRRGTAFCQLELYVEGLQDYEAALK
                     IDPANTVVQNDAEKIRNIIQGTALKSRD"
ORIGIN

  Query Match             67.0%;  Score 846;  DB 10;  Length 1946;
  Best Local Similarity   80.1%;  Pred. No. 1.3e-155;
  Matches 1009;  Conservative    0;  Mismatches  245;  Indels    6;  Gaps    1;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              |||||  | |  || ||||| | ||||||||||||||||  | |  || |||| |||||
Db         48 ATGCCAGTGCGAGTGAGCGAATTCAGCTGGCAGCAGACGCCGGCCACGATCTTCCTGTCG 107

Qy         61 CTGCCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
```

```
                ||||| ||    ||||| |||||  | ||  | ||||| |||||   |||||  || |||
Db        108 CTGCCTCTGCGGGGCGTCTGCGTGCGCGATGCTGACGTATTCTGTGGGGAAAGTTACCTG 167

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||| |||||||||||||||||||||||||    ||||| |||||||||||||| |||| |
Db        168 AAGGTTAACTTTCCTCCATTTTTATTTGAGCTGTTTCTCTATGCTCCCATAGATGATGGG 227

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |  |||||||| |||||||| |||||||| ||| | ||||| |||||||||| ||  |
Db        228 AAGAGCAAAGCCAAGATTGGAAATGACACGATTCTTTTCACATTGTATAAAAAGGAGCCA 287

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |   ||||||| | ||||||| || |||||||||| |||||||||||||| ||||| |||
Db        288 GTTCTGTGGGATAGCCTTTCTATGCCGGGTGTTGATAAAGAGATGATGCAGAGAATAAGA 347

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              |||||||||| || |||||||| |||| |||||||| || ||||||||||||||||
Db        348 GAAAAATCTATCTTGCAAGCACAGGAGAAAGCAAAAGAGGCCACAGAAGCAAAAGCTGTT 407

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              || ||||| ||||| || | |||||||||| | |  |||||||||||||||||||||||||
Db        408 GCCAAGCGAGAAGACCAGAGATACGCACTAGGCGAGATGATGAAGATTGAAGAAGAAGAG 467

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||| ||| || ||||  | ||| |||
Db        468 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGAAAAAGGCAACTAGCGAATTAGAA 527

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              || |||||||||| |||||  |||||||  |  |||||||  | | |||||  | ||||||
Db        528 GCGTGGAAAGAATGTCAAAAGAAAGCTGACGGACAAAAAAGAGTCCAGAGGAAGGAGAAA 587

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
                    |    || |  |  |||| |  |||  |||||  |  | || |
Db        588 ------CCGCCCGAGAGAAAGCAAGCTGAAGAGACCAAAGCTCTAAAACCTCGGGGTTTG 641

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              |  | ||    || | |    ||  |  ||| |||||| |||| ||||| ||||||
Db        642 CCCCGGAAGGCCCCACCCACTCGCCTCCCCACAAGAGGGAGGAATTGGGAAAACATATTT 701

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              |||||||||||||||||||||  | ||||| |||||||| | ||||||||| |||||
Db        702 CCTGAGAAGTTAAAGGAAGACAGAGTCCCTGCGCCTCGCTCCGCTGGCAGTATTCAAATC 761

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              | |||||||||||||| |||||||||| ||||| ||||| |||||||||||||
Db        762 AGCTTTACCCCTCGAGTGTTCCCAACAGCACTTCGGGAATCCCAAGTCGCAGAAGAGGAG 821

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              |||||||| || |||||||| || ||||| ||||| |||| |||||| |  ||||| |
Db        822 GAGTGGCTGCATAAACAAGCAGAAGCACGGAGAGCCATGAGCACTGACCTTCCTGAATTC 881

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              |  || |||||||||||||| | ||| ||||| |||||||| || || ||||||||||||
Db        882 TTTGACTTAAAAGAAGAAGAGAGGAATCCAGACTGGTTGAAAGACAAGGGAAACAAATTG 941

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| |||||||||||||||  | ||||||||||||||||||| |||| ||
Db        942 TTTGCAACAGAAAACTATTTGGCAGCGGTTGATGCATATAATTTAGCCATACGACTGAAC 1001

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              |||||| ||| |||||||||||||| ||||||||||||||||||| ||| |||||||| ||
Db       1002 TGTAAGATCCCATTATTGTATTTGAATCGGGCTGCTTGCCACCTCAAATTAAAAAACCTA 1061

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              |||||||| || || ||||||||||||||| || |||||||  |||||||| | |||
Db       1062 CACAAGGCCATCGAGGACTCTTCTAAGGCACTAGAGTTATTGACACCACCTGTTGCTGAC 1121

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||| |||||||||||||||||| || ||||| || ||||| ||||||||||||||||
Db       1122 AATGCCAATGCAAGAATGAAGGCACACGTCCGACGAGGGACAGCGTTCTGTCAACTAGAA 1181

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              |||||||| |||||| | || |||||||||| |||||||||||||| ||| |||||| |
```

```
Db       1182 TTGTATGTTGAAGGCTTGCAAGATTATGAAGCTGCACTTAAGATTGACCCAGCCAACACA 1241

Qy       1201 ATGTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Db       1242 GTTGTACAGAACGATGCAGAGAAGATTCGGAATATAATTCAAGGGACGGCACTGAAGTCT 1301
```

RESULT 6
BC077575
LOCUS BC077575 2045 bp mRNA linear VRT 12-OCT-2004
DEFINITION Xenopus laevis dyslexia susceptibility 1 candidate 1, mRNA (cDNA clone MGC:83582 IMAGE:5079663), complete cds.
ACCESSION BC077575
VERSION BC077575.1 GI:50603773
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 2045)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 2045)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 3 (bases 1 to 2045)
AUTHORS Klein,S. and Gerhard,D.S.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
REMARK NIH-MGC Project
COMMENT Contact: XGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 153 Row: a Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES Location/Qualifiers
source 1. .2045
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:83582 IMAGE:5079663"
/tissue_type="Oocytes"
/clone_lib="NICHD_XGC_001"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
gene 1. .2045
/gene="dyx1c1-prov"
/note="synonym: MGC83582"
/db_xref="LocusID:446699"
CDS 40. .1305
/gene="dyx1c1-prov"
/codon_start=1
/product="dyx1c1-prov protein"
/protein_id="AAH77575.1"
/db_xref="GI:50603774"
/db_xref="LocusID:446699"
/translation="MPIIANDYTWEQSENDVYISVPLKGTKASASNVLCSESYLKVNF PPFLFEAFLFAPIDDSTSSAKIGNGVILFNLHKKEPAVWDALTLAHMEKEEMQKIREN AVAKEQEKAQEETKAKAERKRENEKYSLEVTMKIDEVERKRIEDVKEEERKKATEELE KWKEKQRIQKLNEMIQKKQIRQENERNKALHTGSLKHCSTATTVQNDRTINVKQKEKI LPPTRSTGCIQIQFTPRVFPTALRESRVAEEEEWLHKQAEARRATNTVDLGLEDLTEE EKNPEWLKDKGNKLFAAGDYLAAVNALNLAIRLNNKIPGLYLNRSACHLKLRNLHKTI EDSSKALELLTPPVPGNASARIKAHVRRGTAFCELELYVEGLQDYEAALKIDPTNQNM KSDAEKIRRVIQSSLPDIERGIMLDIKDL"

ORIGIN

Query Match 38.4%; Score 485.2; DB 5; Length 2045;
Best Local Similarity 63.5%; Pred. No. 5.9e-85;
Matches 789; Conservative 0; Mismatches 423; Indels 30; Gaps 2;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Db         40 ATGCCTATCATAGCGAATGATTATACGTGGGAGCAGAGTGAGAATGATGTTTATATATCC 99

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Db        100 GTCCCCTTAAAAGGTACAAAGGCGTCGGCGTCCAACGTGCTTTGCTCTGAGAGCTATCTA 159

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Db        160 AAGGTGAATTTCCCCCCTTTCCTGTTTGAGGCTTTTCTGTTTGCTCCTATTGATGACAGC 219

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Db        220 ACAAGCAGTGCAAAGATTGGCAATGGAGTTATTCTCTTTAATCTGCACAAGAAGGAGCCT 279

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Db        280 GCCGTATGGGATGCGCTTACCTTGGCACATATGGAAAAAGAAGAAATGCAAAAAATCCGT 339

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Db        340 GAAAATGCTGTTGCAAAAGAACAGGAGAAGGCACAAGAAGAAACAAAAGCCAAAGCAGAA 399

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Db        400 AGGAAGAGAGAAAATGAAAAATATTCTTTAGAAGTCACGATGAAGATCGATGAAGTGGAA 459

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
```

```
Db        460 AGAAAACGTATTGAAGATGTCAAAGAGGAGGAAAGAAAGAAAGCAACAGAGGAACTGGAG 519

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540

Db        520 AAGTGGAAAGAAAAACAGAGGATTCAGAAGTTAAATGAAATGATCCAAAAAAAACA---- 575

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600

Db        576 -------AATAAGACAAGAGAATGAAAGAAATAAAGCATTACACACAGGGTCCTTGAAAC 628

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Db        629 ACTGCTCA------------------ACTGCAACGACTGTACAAAATGACCGTACAATT 669

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Db        670 AACGTCAAGCAGAAAGAAAAGATTTTGCCTCCTACTCGGTCCACTGGGTGCATTCAGATC 729

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Db        730 CAGTTTACACCTCGTGTCTTTCCTACAGCACTCAGAGAATCCCGTGTGGCAGAGGAAGAA 789

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Db        790 GAATGGTTACACAAGCAGGCCGAAGCCCGAAGAGCCACTAATACAGTTGATCTTGGACTT 849

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Db        850 GAAGACTTGACCGAAGAGGAAAAAAATCCAGAGTGGCTTAAAGATAAAGGAAACAAGCTC 909

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Db        910 TTTGCAGCTGGAGACTATCTTGCTGCTGTGAATGCTCTCAACCTGGCAATACGGTTAAAT 969

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Db        970 AATAAGATTCCAGGACTGTATCTAAATCGATCTGCCTGTCATCTTAAACTAAGAAACCTT 1029

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Db       1030 CATAAAACAATTGAGGACTCATCAAAGGCTTTGGAGCTACTTACACCACCCGTTCCAGGC 1089

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Db       1090 AATGCCAGTGCCAGAATTAAGGCTCATGTGAGGCGGGGGACAGCTTTCTGTGAATTGGAA 1149

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Db       1150 CTTTATGTAGAAGGTCTCCAAGACTACGAAGCTGCCCTTAAAATTGATCCCACCAATCAA 1209

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAA 1242

Db       1210 AACATGAAATCTGATGCTGAAAAGATTCGACGGGTCATCCAA 1251


RESULT 7
AR413597
LOCUS       AR413597                489 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 1234 from patent US 6639063.
ACCESSION   AR413597
VERSION     AR413597.1  GI:40168707
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
  TITLE     EST's and encoded human proteins
  JOURNAL   Patent: US 6639063-A 1234 28-OCT-2003;
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           34.3%;  Score 432.6;  DB 6;  Length 489;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-74;
  Matches  432;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Db         57 ATGCCTCTTCAGGTTAGCGATTACAKCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 116

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Db        117 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 176

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Db        177 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 236

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Db        237 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 296

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Db        297 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 356

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Db        357 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 416

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Db        417 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 476

Qy        421 AGGAAAAAAAATAG 433

Db        477 AGGAAAAAAAATAG 489


RESULT 8
AX970431
LOCUS       AX970431                489 bp    DNA     linear   PAT 15-JAN-2004
DEFINITION  Sequence 1234 from Patent EP1104808.
ACCESSION   AX970431
VERSION     AX970431.1  GI:40977781
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
  TITLE     ESTs and encoded human proteins
  JOURNAL   Patent: EP 1104808-A 1234 06-JUN-2001;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             57. .>488
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAF15539.1"
                     /db_xref="GI:40977782"
                     /translation="MPLQVSDYXWQQTKTAVFLSLPLKGVCVRDTDVFCTENYLKVNF
                     PPFLFEAFLYAPIDDESSKAKIGNDTIVFTLYKKEAAMWETLSVTGVDKEMMQRIREK
                     SILQAQERAKEATEAKAAAKREDQKYALSVMMKIEEEERKKI"
ORIGIN

  Query Match           34.3%;  Score 432.6;  DB 6;  Length 489;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-74;
  Matches  432;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db         57 ATGCCTCTTCAGGTTAGCGATTACAKCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 116

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 176

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 236

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 296

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 356

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 416

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 476

Qy        421 AGGAAAAAAAATAG 433
              ||||||||||||||
Db        477 AGGAAAAAAAATAG 489
```

```
RESULT 9
BD109150
LOCUS       BD109150                489 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  EST and encoded human protein.
ACCESSION   BD109150
VERSION     BD109150.1  GI:23203968
KEYWORDS    JP 2002010789-A/1227.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
  TITLE     EST and encoded human protein
  JOURNAL   Patent: JP 2002010789-A 1227 15-JAN-2002;
            GENSET CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002010789-A/1227
            PD   15-JAN-2002
            PF   07-AUG-2000 JP 2000280989
            PR   05-AUG-1999 US    60/147499
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,SEVELIN JOBERT,JEAN EVE  PI
             GIORDANO
            PC   C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC
             C12N1/21,
            PC   C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC
            C12N15/00
            CC   EST and encoded human protein
            FH   Key             Location/Qualifiers
            FT   CDS             57. .488.
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           34.3%;  Score 432.6;  DB 6;  Length 489;
  Best Local Similarity  99.8%;  Pred. No. 1.3e-74;
  Matches  432;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||:||||||||||||||||||||||||||||||||||||
Db         57 ATGCCTCTTCAGGTTAGCGATTACAKCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 116

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 176

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 236

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 296

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 356

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 416

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 476

Qy        421 AGGAAAAAAAATAG 433
              ||||||||||||||
Db        477 AGGAAAAAAAATAG 489
```

```
RESULT 10
CQ685075
LOCUS       CQ685075                464 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 30001 from Patent WO02070737.
ACCESSION   CQ685075
VERSION     CQ685075.1  GI:42211607
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liew,C.C., Marshall,W.E. and Zhang,H.
  TITLE     Compositions and methods relating to osteoarthritis
  JOURNAL   Patent: WO 02070737-A 30001 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           31.8%;  Score 401.6;  DB 6;  Length 464;
  Best Local Similarity  98.3%;  Pred. No. 1.5e-68;
  Matches  404;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 108

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 168

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 228
```

```
Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||| |||||||||||||||||| ||||
Db        229 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGNCTTCACCTTGTATAAAAAAGGAGCG 288

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |
Db        289 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTNGACAAAGAGATGATGCAAAGAATTANA 348

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 408

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAA 411
              |||||||||||||||||||||||||||||||||||||||||||||    |||
Db        409 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGGGAGAA 459
```

RESULT 11
BC065881
LOCUS BC065881 1443 bp mRNA linear VRT 30-JUN-2004
DEFINITION Danio rerio zgc:77853, mRNA (cDNA clone MGC:77853 IMAGE:7002321), complete cds.
ACCESSION BC065881
VERSION BC065881.1 GI:41389019
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 1443)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 1443)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

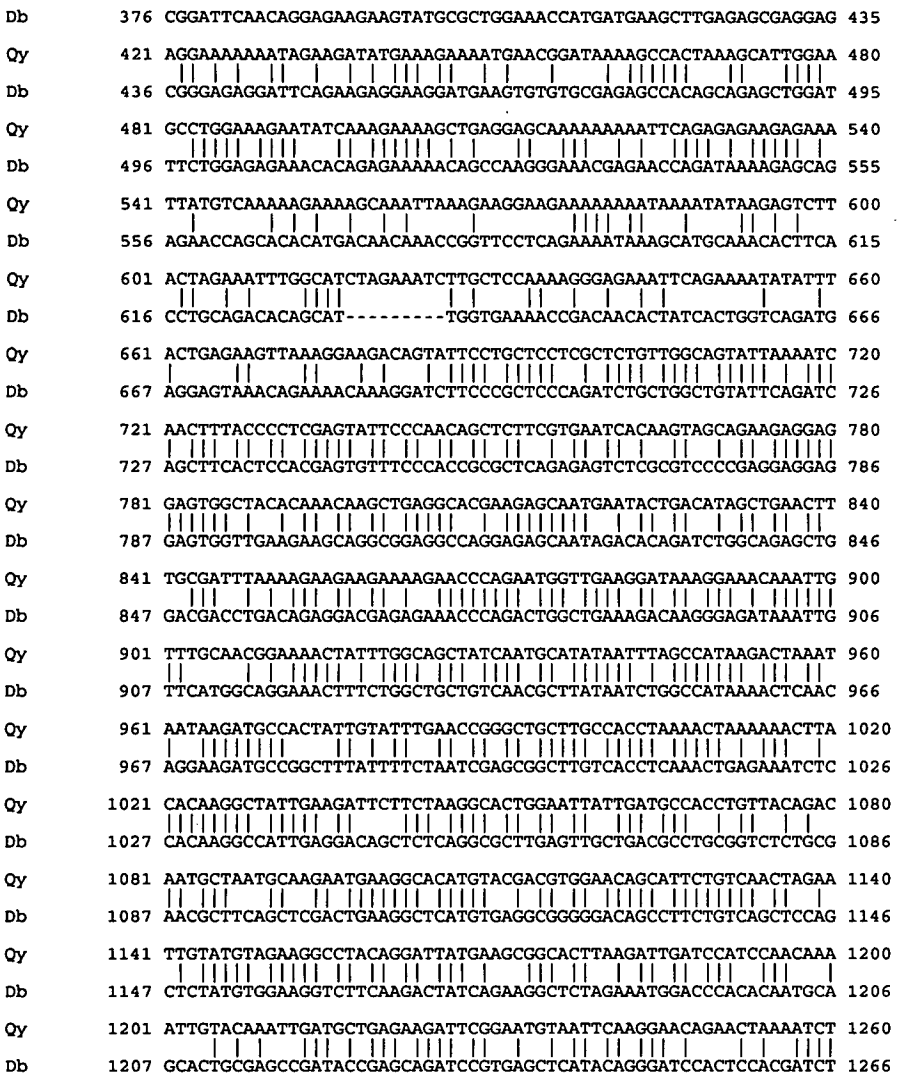
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 147 Row: p Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES Location/Qualifiers
source 1. .1443
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:77853 IMAGE:7002321"
/tissue_type="Whole, adult zebrafish"
/clone_lib="NIH_ZGC_10"
/lab_host="DH10B"
/note="Vector: pExpress1"
gene 1. .1443
/gene="zgc:77853"
/note="synonym: MGC77853"
/db_xref="LocusID:402988"
CDS 19. .1281
/gene="zgc:77853"
/codon_start=1
/product="hypothetical protein zgc:77853"
/protein_id="AAH65881.1"
/db_xref="GI:41389020"
/db_xref="LocusID:402988"
/translation="MPLIVRDHTWTQNLSTVYISVPLKAVKTANVHIICTDDYLKVSF
PPFLFEVFLFASINEEKSEARIGNGAAVFTLQKRRDEMWEQLCTNIDKEKQKQIREQA
VLKVQEKAVEKSKTKATRIQQEKKYALETMMKLESEERERIQKRKDEVCARATAELDF
WRETQRKTAKGNENQIKEQRTSTHDNKPVPQKIKHANTSPADTALVKTDNTITGQMRS
KQKTKDLPAPRSAGCIQISFTPRVFPTALRESRVPEEEEWLKKQAEARRAIDTDLAEL
DDLTEDERNPDWLKDKGDKLFMAGNFLAAVNAYNLAIKLNRKMPALFSNRAACHLKLR
NLHKAIEDSSQALELLTPAVSANASARLKAHVRRGTAFCQLQLYVEGLQDYQKALEMD
PHNAALRADTEQIRELIQGSTPRSKDEN"

ORIGIN

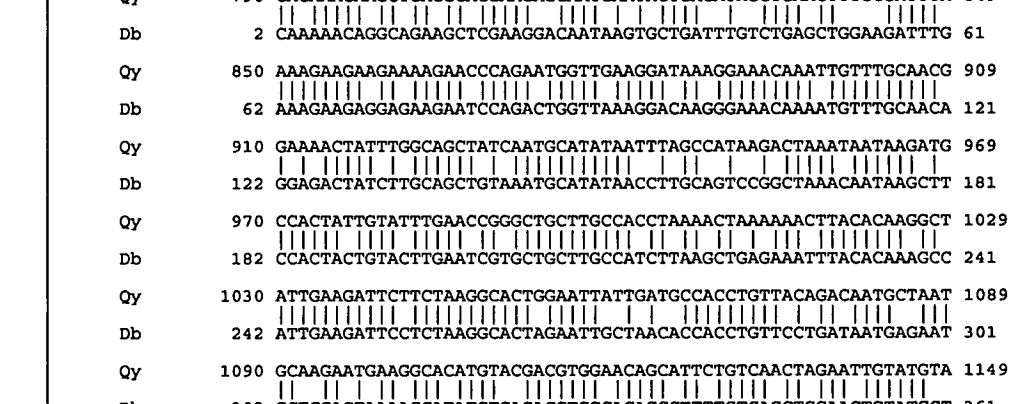
Query Match 30.9%; Score 390.2; DB 5; Length 1443;
Best Local Similarity 58.0%; Pred. No. 2.4e-66;
Matches 733; Conservative 0; Mismatches 518; Indels 12; Gaps 2;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| ||    || || ||| |||  |||   || |    ||   | || |   |
Db         19 ATGCCGCTGATAGTAAGAGATCACACATGGACACAAAACCTGAGCACAGTTTACATCAGC 78

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
                | ||  | ||||  |||      |  | |      || |  ||||||||||  |||| |||
Db         79 GTCCCTTTAAAAGCAGTGAAAACCGCAAATGTACACATCATCTGCACGGACGACTACCTG 138

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              |||||||  || || || ||| | || ||||  ||  | |  ||  | ||| | || |||
Db        139 AAGGTCAGTTTCCCACCGTTTCTCTTCGAGGTTTTCTTATTCGCCTCAATAAATGAGGAG 198

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |  ||  |||| | |||||| || |    |    || |||||  |  | || |    ||
Db        199 AAAAGTGAAGCCAGGATTGGAAACGGTGCTGCAGTTTTCACTCTTCAGAAGAGGAGAGAT 258

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |  |||||||||    || | |   ||      | || |||||||     | |   |||| ||
Db        259 GAAATGTGGGAGCAGCTCTGT---ACAAACATCGATAAAGAGAAACAGAAGCAAATTCGA 315

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||| || |  || |  |||  || |||| ||||   |||    |  || | |||||  |
Db        316 GAACAAGCCGTTCTCAAAGTTCAGGAGAAAGCAGTGGAAAAGTCTAAAACCAAAGCCACC 375

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
                |  |  | ||  | || || || ||      ||||||||| ||||   || |||
```



RESULT 12
CR387522
LOCUS CR387522 761 bp mRNA linear VRT 05-APR-2004
DEFINITION Gallus gallus finished cDNA, clone ChEST505n19.
ACCESSION CR387522
VERSION CR387522.1 GI:46240281
KEYWORDS .
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 761)
AUTHORS Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E., Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V., Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R., Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G., Tickle,C. and Wilson,S.A.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA sequencing project.
This sequence is from the BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection, from a library constructed by Elizabeth Bosch. cDNA was prepared from RNA extracted from hearts, normalised, and poly A-trimmed. EcoRI-NotI cut cDNA was then ligated into the vector. Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia coli DH10B.
FEATURES Location/Qualifiers
source 1. .761
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 15I"
/db_xref="taxon:9031"
/clone="ChEST505n19"
/clone_lib="CSEQCHN60"
/dev_stage="adult"
ORIGIN

Query Match 23.6%; Score 298; DB 5; Length 761;
Best Local Similarity 76.8%; Pred. No. 2.8e-48;
Matches 364; Conservative 0; Mismatches 110; Indels 0; Gaps 0;



RESULT 13
CR389904
LOCUS CR389904 1133 bp mRNA linear VRT 20-APR-2004
DEFINITION Gallus gallus finished cDNA, clone ChEST271n13.
ACCESSION CR389904

```
VERSION     CR389904.1  GI:46428549
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1133)
  AUTHORS   Boardman,P.E., Bonfield,J.K., Brown,W.R.A., Carder,C., Chalk,S.E.,
            Croning,M.D.R., Davies,R.M., Francis,M.D., Grafham,D.V.,
            Hubbard,S.J., Humphray,S.J., Hunt,P.J., Maddison,M., McLaren,S.R.,
            Niblett,D., Overton,I.M., Rogers,J., Scott,C.E., Taylor,R.G.,
            Tickle,C. and Wilson,S.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2004) Sanger Institute, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: chickest@bms.umist.ac.uk
COMMENT     BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
            sequencing project.
            This sequence is from the
            BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
            from a library constructed by Elizabeth Bosch. cDNA was prepared
            from RNA extracted from ovary, normalised, and poly A-trimmed.
            EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
            pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
            coli DH10B.
FEATURES             Location/Qualifiers
     source          1. .1133
                     /organism="Gallus gallus"
                     /mol_type="mRNA"
                     /strain="Layer"
                     /db_xref="taxon:9031"
                     /clone="ChEST271n13"
                     /clone_lib="CSEQRBN19"
                     /dev_stage="adult"
ORIGIN

  Query Match             21.8%;  Score 275;  DB 5;  Length 1133;
  Best Local Similarity   66.4%;  Pred. No. 8.8e-44;
  Matches  395;  Conservative    0;  Mismatches  200;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| ||   | |  | || |||||||||| |||||      | || || |  || ||
Db         73 ATGCCGCTGTGGCTGCGGGAGTACAGCTGGCGGCAGAGCGGCGCCGCCGTGTACCTCTCG 132

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              |||||| |     ||| || | ||||  | | |  || | |||||||| ||    || |||
Db        133 CTGCCCGTGCGCGGCCTGAGGGTCACCGCCGCCAACATCTTCTGCACCGAGCGGTACCTG 192

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||| | | | ||||| |||||||| || ||  |||| |||||||| || || ||
Db        193 AAGGTAAGCGTCCCTCCCTTTTTATTCGAAGCTGTTCTGTATGCTCCTATTGATGAGACA 252

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |  |||| ||||||||| || ||||    | || |||||||  ||||||||||||||||
Db        253 AATAGCACAGCAAAGATCGGCAATGGAGTCGTTTTCTTCACTCTGTATAAAAAAGAAGCG 312

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |||||||||||| |||| |||| |  |||| |||| |||    |||||  |  | |||
Db        313 GCCATGTGGGAGTCCCTGGCTGTAGCAAATGTTAACAAGGAGGAACTGCAACGTCTAAGA 372

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              || ||  || || |  ||||||| || |  |||||||  |  ||||||||||| |
Db        373 GAGAATGCTGTTCTGAAAGCACACGAAAAGGCAAAAGAGGAGACAGAAGCAAAAAAAGTT 432

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              |||| | |||| || | || || ||  |     |  |  |||||| | |||||| |||
Db        433 ACAAAACAGGAACATAAGAAGTATGCTTTGGAGGCTACAATGAAGCTAGAAGAAGCAGAA 492

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              || |||| ||| |||||| |||||||| | ||  |    |||| | |||| |  |||||
Db        493 AGAAAAAGAATTGAAGATCTGAAAGAACAGGAGAGACAGAAAGTCGCTAAGGAGTTGGAG 552

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
```

```
                |||| |    | ||||   | | |||| | ||||| |   | || |   ||| | ||
Db        553 CTATGGAGAAGTCAGCAAAACGATGTTGAGAAACAAAAGAGGGTACAAAAGAAAGGGGAA 612

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGA 595
                ||  | || ||| | ||||| | || ||    | || | ||| |||| | || |
Db        613 CTACATGAAGAAGTAGAGCAACTAAAGGAGAAGAAAAGAGAAAAAAAAAACAAAA 667
```

```
RESULT 14
AC022083/c
LOCUS       AC022083              114149 bp    DNA     linear   PRI 12-MAY-2001
DEFINITION  Homo sapiens chromosome 15 clone CTD-2137J4 map 15q21.3, complete
            sequence.
ACCESSION   AC022083
VERSION     AC022083.6  GI:14029048
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 114149)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 114149)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2000) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 114149)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-MAY-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On May 12, 2001 this sequence version replaced gi:13621222.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399

            Note: Data from overlapping clones AC013355 [Drafting center:
            UWMSC], AC012674 [Drafting center: BCM] and AC012378 [Drafting
            center: UWMSC] were added for finishing.
FEATURES             Location/Qualifiers
     source          1. .114149
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21.3"
                     /clone="CTD-2137J4"
                     /clone_lib="Cal Tech human BAC library D"
                     /note="This clone overlaps RP11-178D12 AC013355,
                     RP11-458H3 AC012674 and RP11-420M1 AC012378. Data from
                     overlapping BACs were added and the consensus sequence
                     determined from CTD-2137J4 to the extent possible."
     misc_feature    1. .47896
                     /note="oerlap with RP11-178D12, AC013355"
     misc_feature    18530. .18640
```

```
                     /note="sequence data generated from subcloned PCR product"
     unsure          54340. .54368
                     /note="low quality data"
     unsure          90483. .90630
                     /note="low quality data"
     misc_feature    95860. .114149
                     /note="overlap with RP11-458H3, AC012674"
     misc_feature    102476. .114149
                     /note="overlap with RP11-420M1, AC012378"
ORIGIN

  Query Match             19.1%;  Score 241;  DB 9;  Length 114149;
  Best Local Similarity   92.7%;  Pred. No. 3.2e-37;
  Matches  253;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy        398 TGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGA 457
              || |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31582 TGTTATAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGA 31523

Qy        458 TAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAA 517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31522 TAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAA 31463

Qy        518 AAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAA 577
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31462 AAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAA 31403

Qy        578 AAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAG 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      31402 AAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAG 31343

Qy        638 GGAGAAATTCAGAAAATATATTTACTGAGAAGT 670
              | | ||      | ||| | || | |    |||
Db      31342 GTAAAATGATTAAGAATTTCTTGAGTAGTGAGT 31310


RESULT 15
AC021668/c
LOCUS       AC021668               43886 bp    DNA     linear   HTG 29-MAR-2001
DEFINITION  Homo sapiens chromosome 15 clone RP11-133O6 map 15, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC021668
VERSION     AC021668.4  GI:13488020
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43886)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 15, clone RP11-133O6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 43886)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Mar 29, 2001 this sequence version replaced gi:11612359.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L1276
                Center clone name: 133_O_6
            --------------.
            * NOTE: This record contains 53 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      745: contig of 745 bp in length
            *      746      845: gap of 100 bp
            *      846     1576: contig of 731 bp in length
            *     1577     1676: gap of 100 bp
            *     1677     2410: contig of 734 bp in length
            *     2411     2510: gap of 100 bp
            *     2511     3232: contig of 722 bp in length
            *     3233     3332: gap of 100 bp
            *     3333     4060: contig of 728 bp in length
            *     4061     4160: gap of 100 bp
            *     4161     4891: contig of 731 bp in length
            *     4892     4991: gap of 100 bp
            *     4992     5708: contig of 717 bp in length
            *     5709     5808: gap of 100 bp
            *     5809     6537: contig of 729 bp in length
            *     6538     6637: gap of 100 bp
            *     6638     7344: contig of 707 bp in length
            *     7345     7444: gap of 100 bp
            *     7445     8167: contig of 723 bp in length
            *     8168     8267: gap of 100 bp
            *     8268     8995: contig of 728 bp in length
            *     8996     9095: gap of 100 bp
            *     9096     9839: contig of 744 bp in length
            *     9840     9939: gap of 100 bp
            *     9940    10653: contig of 714 bp in length
            *    10654    10753: gap of 100 bp
            *    10754    11497: contig of 744 bp in length
            *    11498    11597: gap of 100 bp
            *    11598    12315: contig of 718 bp in length
            *    12316    12415: gap of 100 bp
            *    12416    13136: contig of 721 bp in length
            *    13137    13236: gap of 100 bp
            *    13237    13967: contig of 731 bp in length
            *    13968    14067: gap of 100 bp
            *    14068    14802: contig of 735 bp in length
            *    14803    14902: gap of 100 bp
            *    14903    15642: contig of 740 bp in length
            *    15643    15742: gap of 100 bp
            *    15743    16477: contig of 735 bp in length
            *    16478    16577: gap of 100 bp
            *    16578    17301: contig of 724 bp in length
            *    17302    17401: gap of 100 bp
            *    17402    18129: contig of 728 bp in length
            *    18130    18229: gap of 100 bp
            *    18230    18958: contig of 729 bp in length
            *    18959    19058: gap of 100 bp
            *    19059    19787: contig of 729 bp in length
            *    19788    19887: gap of 100 bp
            *    19888    20627: contig of 740 bp in length
```

```
            *    20628    20727: gap of 100 bp
            *    20728    21449: contig of 722 bp in length
            *    21450    21549: gap of 100 bp
            *    21550    22275: contig of 726 bp in length
            *    22276    22375: gap of 100 bp
            *    22376    23103: contig of 728 bp in length
            *    23104    23203: gap of 100 bp
            *    23204    23939: contig of 736 bp in length
            *    23940    24039: gap of 100 bp
            *    24040    24775: contig of 736 bp in length
            *    24776    24875: gap of 100 bp
            *    24876    25602: contig of 727 bp in length
            *    25603    25702: gap of 100 bp
            *    25703    26418: contig of 716 bp in length
            *    26419    26518: gap of 100 bp
            *    26519    27260: contig of 742 bp in length
            *    27261    27360: gap of 100 bp
            *    27361    28107: contig of 747 bp in length
            *    28108    28207: gap of 100 bp
            *    28208    28931: contig of 724 bp in length
            *    28932    29031: gap of 100 bp
            *    29032    29776: contig of 745 bp in length
            *    29777    29876: gap of 100 bp
            *    29877    30592: contig of 716 bp in length
            *    30593    30692: gap of 100 bp
            *    30693    31422: contig of 730 bp in length
            *    31423    31522: gap of 100 bp
            *    31523    32255: contig of 733 bp in length
            *    32256    32355: gap of 100 bp
            *    32356    33073: contig of 718 bp in length
            *    33074    33173: gap of 100 bp
            *    33174    33891: contig of 718 bp in length
            *    33892    33991: gap of 100 bp
            *    33992    34711: contig of 720 bp in length
            *    34712    34811: gap of 100 bp
            *    34812    35546: contig of 735 bp in length
            *    35547    35646: gap of 100 bp
            *    35647    36407: contig of 761 bp in length
            *    36408    36507: gap of 100 bp
            *    36508    37258: contig of 751 bp in length
            *    37259    37358: gap of 100 bp
            *    37359    38073: contig of 715 bp in length
            *    38074    38173: gap of 100 bp
            *    38174    38908: contig of 735 bp in length
            *    38909    39008: gap of 100 bp
            *    39009    39729: contig of 721 bp in length
            *    39730    39829: gap of 100 bp
            *    39830    40548: contig of 719 bp in length
            *    40549    40648: gap of 100 bp
            *    40649    41383: contig of 735 bp in length
            *    41384    41483: gap of 100 bp
            *    41484    42222: contig of 739 bp in length
            *    42223    42322: gap of 100 bp
            *    42323    43055: contig of 733 bp in length
            *    43056    43155: gap of 100 bp
            *    43156    43886: contig of 731 bp in length.
FEATURES             Location/Qualifiers
     source          1. .43886
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15"
                     /clone="RP11-133O6"
                     /clone_lib="RPCI-11 Human Male BAC"
ORIGIN

  Query Match             19.0%;  Score 239.4;  DB 2;  Length 43886;
  Best Local Similarity   92.3%;  Pred. No. 6.9e-37;
  Matches  252;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy        398 TGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGA 457
              || |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       6300 TGTTATAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGA 6241

Qy        458 TAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAA 517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       6240 TAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAA 6181

Qy        518 AAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAA 577
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||

Db       6180 AAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAA 6121

Qy        578 AAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAG 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       6120 AAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAG 6061

Qy        638 GGAGAAATTCAGAAAATATATTTACTGAGAAGT 670
              | | ||      | ||| | || | |    |||

Db       6060 GTAAAATGATTAAGAATTTCTTGAGTAGTGAGT 6028


Search completed: July  8, 2005, 03:29:22
Job time : 5662 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:36:17 ; Search time 755 Seconds
(without alignments)
9902.826 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1263	100.0	1263	9	ADB16964	Adb16964 Human DYX
	2	1263	100.0	1993	9	ADB16965	Adb16965 cDNA sequ
	3	1253.4	99.2	1263	9	ADB16939	Adb16939 Pygmy chi
	4	1251.8	99.1	1263	9	ADB16933	Adb16933 Chimpanze
	5	1248.6	98.9	1263	9	ADB16935	Adb16935 Gorilla D
	6	1243.8	98.5	1263	9	ADB16937	Adb16937 Orangutan
	7	1049	83.1	1641	10	ADC30210	Adc30210 Human nov
	8	1047.4	82.9	1559	11	ADM01890	Adm01890 Human cDN
	9	840.8	66.6	1697	9	ADB16924	Adb16924 cDNA sequ
	10	524	41.5	608	10	ADC32116	Adc32116 Human nov
	11	369.8	29.3	488	9	ACH35463	Ach35463 Human end
	12	336	26.6	458	9	ACH23091	Ach23091 Human adu
	13	245.2	19.4	1383	5	AAS70018	Aas70018 DNA encod
	14	239.4	19.0	49806	9	ADB16927	Adb16927 Human DYX
	15	178.8	14.2	313	3	AAC26799	Aac26799 Human sec
	16	162.8	12.9	164	3	AAC30498	Aac30498 Human sec
	17	156.6	12.4	49939	9	ADB16928	Adb16928 Human DYX
c	18	97.6	7.7	14006	6	ABL33958	Abl33958 Human imm
c	19	96.8	7.7	16033	6	ABL33404	Abl33404 Human imm
c	20	96.4	7.6	700	10	ACD92384	Acd92384 Human col
	21	94.6	7.5	6644	2	AAX33181	Aax33181 Base sequ
	22	94.6	7.5	7372	2	AAX33182	Aax33182 Base sequ
	23	94.6	7.5	7797	2	AAX33180	Aax33180 Cowpox vi
	24	94.6	7.5	7996	2	AAX33184	Aax33184 Base sequ
	25	93.4	7.4	447	13	ACN54792	Acn54792 Cotton an
c	26	93.2	7.4	529	8	ABX52180	Abx52180 Bovine ES
	27	93.2	7.4	2520	12	ADJ48201	Adj48201 Maize oil
	28	92.8	7.3	627	13	ACN54555	Acn54555 Cotton an
c	29	92.8	7.3	6668	6	ABL33697	Abl33697 Human imm
c	30	92.8	7.3	17934	6	ABL33719	Abl33719 Human imm
	31	92.2	7.3	556	6	ABQ36997	Abq36997 Oligonucl
c	32	92.2	7.3	556	6	ABQ36996	Abq36996 Oligonucl
c	33	92.2	7.3	16633	6	ABN79984	Abn79984 Human che
	34	91.4	7.2	1000	12	ADQ62833	Adq62833 Homopoly-
c	35	91.4	7.2	1000	12	ADQ62832	Adq62832 Homopoly-
	36	91.4	7.2	1300	12	ADP85917	Adp85917 Synthetic
	37	91.2	7.2	3505	12	ADJ48200	Adj48200 Maize oil
	38	89.8	7.1	2226	12	ADI43251	Adi43251 Plant tra
	39	89.8	7.1	2226	12	ADO02941	Ado02941 Soybean o
c	40	89.6	7.1	34769	4	AAS46774	Aas46774 Tumour su
c	41	89.4	7.1	113515	6	ABL34174	Abl34174 Human imm
	42	88.4	7.0	537	13	ACN53339	Acn53339 Cotton an
c	43	88.2	7.0	1179	6	ABQ49604	Abq49604 Oligonucl
	44	88.2	7.0	1179	6	ABQ49605	Abq49605 Oligonucl
	45	87.6	6.9	529	13	ACN56029	Acn56029 Cotton an

ALIGNMENTS

RESULT 1
ADB16964
ID ADB16964 standard; cDNA; 1263 BP.
XX
AC ADB16964;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human DYXC1 cDNA with single nucleotide polymorphisms.
XX
KW gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW chromosome 15q21; reading disability; phonological processing;
KW rapid naming; verbal short-term memory; single nucleotide polymorphism;
KW SNP.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1. .1263
FT /*tag= a
FT /product= "DYXC1 protein"
FT variation replace(4,t)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(572,a)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(1249,t)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(1259,g)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.

XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
DR P-PSDB; ADB16923.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 1; Page 47; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide
CC represents the coding sequence of human DYXC1 cDNA of the invention.
XX
SQ Sequence 1263 BP; 491 A; 214 C; 275 G; 283 T; 0 U; 0 Other;

Query Match 100.0%; Score 1263; DB 9; Length 1263;
Best Local Similarity 100.0%; Pred. No. 1.3e-263;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263
```

RESULT 2
ADB16965
ID ADB16965 standard; cDNA; 1993 BP.
XX
AC ADB16965;
XX
DT 20-NOV-2003 (first entry)
XX
DE cDNA sequence of the human DYXC1 mRNA.
XX
KW gene; ss; human; DYXC1; dyslexia; neurological disorder;
KW chromosome 15q21; reading disability; phonological processing;
KW rapid naming; verbal short-term memory; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	variation	replace(205,t)
FT		/*tag= a
FT		/standard_name= "Single nucleotide polymorphism"
FT	CDS	369. .1631
FT		/*tag= b
FT		/product= "DYXC1 protein"

XX

PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.
XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
DR P-PSDB; ADB16923.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 10; Page 48-50; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC cDNA sequence of the wild type human DYXC1 mRNA of the invention.
XX
SQ Sequence 1993 BP; 693 A; 368 C; 430 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 1263; DB 9; Length 1993;
Best Local Similarity 100.0%; Pred. No. 1.4e-263;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631
```

RESULT 3
ADB16939
ID ADB16939 standard; cDNA; 1263 BP.
XX
AC ADB16939;
XX
DT 20-NOV-2003 (first entry)
XX
DE Pygmy chimpanzee DYXC1 cDNA sequence.
XX
KW gene; ss; pygmy chimpanzee; DYXC1; dyslexia; neurological disorder;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Pan paniscus.
XX

```
FH   Key             Location/Qualifiers
FT   CDS             1. .1263
FT                   /*tag= a
FT                   /product= "DYXC1 protein"
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16940.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 29; Page 126-128; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide
CC   sequence is the pygmy chimpanzee DYXC1 cDNA homologous to the human DYXC1
CC   gene of the invention.
XX
SQ   Sequence 1263 BP; 493 A; 213 C; 274 G; 283 T; 0 U; 0 Other;

  Query Match             99.2%;  Score 1253.4;  DB 9;  Length 1263;
  Best Local Similarity   99.5%;  Pred. No. 1.5e-261;
  Matches 1257;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTGCACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263


RESULT 4
ADB16933
ID   ADB16933 standard; cDNA; 1263 BP.
XX
AC   ADB16933;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Chimpanzee DYXC1 cDNA sequence.
XX
```

KW gene; ss; chimpanzee; DYXC1; dyslexia; neurological disorder;
KW reading disability; phonological processing; rapid naming;
KW verbal short-term memory.
XX
OS Pan troglodytes.
XX
FH Key Location/Qualifiers
FT CDS 1. .1263
FT /*tag= a
FT /product= "DYXC1 protein"
XX
PN WO2003068814-A1.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-FI000110.
XX
PR 12-FEB-2002; 2002US-0355782P.
XX
PA (LICN) LICENTIA LTD.
XX
PI Kere J, Taipale M, Nopola-Hemmi J, Kaminen N;
XX
DR WPI; 2003-646482/61.
DR P-PSDB; ADB16934.
XX
PT New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT e.g. reading, phonological processing, rapid naming or verbal short-term
PT memory, or for diagnosing dyslexia or assessing the predisposition to
PT dyslexia.
XX
PS Claim 29; Page 116-118; 135pp; English.
XX
CC This invention relates to a novel isolated human gene DYXC1 that is
CC functionally related to dyslexia, more particularly it describes single
CC nucleotide polymorphisms thought to predispose an individual in to
CC developing dyslexia. This is a neurological disorder with a genetic basis
CC (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC a specific reading disability. Specifically, DYXC1 is can be useful in
CC study of brain processes such as reading, phonological processing, rapid
CC naming and verbal short-term memory. Accordingly, the present invention
CC describes methods and materials for analysing allelic variations in the
CC DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC antibodies used in the diagnosis of dyslexia. This polynucleotide
CC sequence is the chimpanzee DYXC1 cDNA homologous to the human DYXC1 gene
CC of the invention.
XX
SQ Sequence 1263 BP; 494 A; 213 C; 272 G; 284 T; 0 U; 0 Other;

Query Match 99.1%; Score 1251.8; DB 9; Length 1263;
Best Local Similarity 99.4%; Pred. No. 3.4e-261;
Matches 1256; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCCGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCATGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263
```

RESULT 5
ADB16935
ID ADB16935 standard; cDNA; 1263 BP.
XX

```
AC   ADB16935;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Gorilla DYXC1 cDNA sequence.
XX
KW   gene; ss; gorilla; DYXC1; dyslexia; neurological disorder;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory.
XX
OS   Gorilla gorilla.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1263
FT                   /*tag= a
FT                   /product= "DYXC1 protein"
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16936.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 29; Page 119-121; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide
CC   sequence is the gorilla DYXC1 cDNA homologous to the human DYXC1 gene of
CC   the invention.
XX
SQ   Sequence 1263 BP; 491 A; 212 C; 275 G; 285 T; 0 U; 0 Other;

  Query Match             98.9%;  Score 1248.6;  DB 9;  Length 1263;
  Best Local Similarity   99.3%;  Pred. No. 1.7e-260;
  Matches 1254;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGTGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAGAAATTCAGAGAGAAGAGAAG 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATTAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263
```

```
RESULT 6
ADB16937
ID   ADB16937 standard; cDNA; 1263 BP.
XX
AC   ADB16937;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Orangutan DYXC1 cDNA sequence.
XX
KW   gene; ss; orangutan; DYXC1; dyslexia; neurological disorder;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory.
XX
OS   Pongo pygmaeus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1263
FT                   /*tag= a
FT                   /product= "DYXC1 protein"
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16938.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 29; Page 123-124; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide
CC   sequence is the orangutan DYXC1 cDNA homologous to the human DYXC1 gene
CC   of the invention.
XX
SQ   Sequence 1263 BP; 493 A; 218 C; 272 G; 280 T; 0 U; 0 Other;

  Query Match             98.5%;  Score 1243.8;  DB 9;  Length 1263;
  Best Local Similarity   99.0%;  Pred. No. 1.8e-259;
  Matches 1251;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGACGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCACAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||| |||||||||| ||||||| |||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATTAAAATACAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              |||||||||| |||||||||||| |||||||||||||| |||||||||||||||||||||
Db        601 ACTAGAAATTCGGCATCTAGAAACCTTGCTCCAAAAGGAAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
                |||
Db       1261 TAA 1263
```

RESULT 7
ADC30210
ID ADC30210 standard; cDNA; 1641 BP.
XX
AC ADC30210;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:292.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 15q21.3; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31181.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 292; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of peventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1641 BP; 614 A; 272 C; 331 G; 424 T; 0 U; 0 Other;

Query Match 83.1%; Score 1049; DB 10; Length 1641;
Best Local Similarity 100.0%; Pred. No. 2.9e-217;
Matches 1049; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 163

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 223

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 283

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 343

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 403

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 463

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 523

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 583

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 643

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 703

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 763

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 823

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 883

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        884 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 943

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1003

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1063

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1123

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1124 CACAAGGCTATTGAAGATTCTTCTAAGGC 1152
```

RESULT 8
ADM01890
ID ADM01890 standard; cDNA; 1559 BP.
XX
AC ADM01890;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human cDNA of the invention SEQ ID NO:575.
XX
KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
PN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-723558/69.
DR P-PSDB; ADM04333.
XX
PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
PS Claim 1; SEQ ID NO 575; 305pp; English.
XX
CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.
XX
SQ Sequence 1559 BP; 592 A; 252 C; 309 G; 406 T; 0 U; 0 Other;

Query Match 82.9%; Score 1047.4; DB 11; Length 1559;
Best Local Similarity 99.9%; Pred. No. 6.4e-217;
Matches 1048; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy           1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy          61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162

Qy         121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy         181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy         241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy         301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy         361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy         421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         463 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy         481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy         541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTT 642

Qy         601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy         661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy         721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy         781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy         841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy         901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy         961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy        1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
               |||||||||||||||||||||||||||||
Db        1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091
```

```
RESULT 9
ADB16924
ID   ADB16924 standard; cDNA; 1697 BP.
XX
AC   ADB16924;
XX
DT   20-NOV-2003  (first entry)
XX
DE   cDNA sequence of the murine DYXC1 mRNA.
XX
KW   gene; ss; mouse; DYXC1; dyslexia; neurological disorder;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory; murine.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             48. .1310
FT                   /*tag= a
FT                   /product= "DYXC1 protein"
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
DR   P-PSDB; ADB16925.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Disclosure; Page 51-53; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC   cDNA sequence of the murine DYXC1 mRNA of the invention.
XX
SQ   Sequence 1697 BP; 534 A; 360 C; 441 G; 362 T; 0 U; 0 Other;

  Query Match             66.6%;  Score 840.8;  DB 9;  Length 1697;
  Best Local Similarity   80.3%;  Pred. No. 3.5e-172;
  Matches 1012;  Conservative    0;  Mismatches  242;  Indels    6;  Gaps    2;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              |||||  | |  || ||||| | |||||||||||||||||  | |  || |||| |||||
Db         48 ATGCCAGTGCGAGTGAGCGAATTCAGCTGGCAGCAGACGCCGGCCACGATCTTCCTGTCG 107

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||| ||    ||||| |||||  | ||  | ||||| |||||   |||||  || |||
Db        108 CTGCCTCTGCGGGGCGTCTGCGTGCGCGATGCTGACGTATTCTGTGGGGAAAGTTACCTG 167

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||| |||||||||||||||||||||||    ||||| |||||||||||||| |||| |
Db        168 AAGGTTAACTTTCCTCCATTTTTATTTGAGCTGTTTCTCTATGCTCCCATAGATGATGGG 227

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |  |||||||| |||||||| |||||||| ||| | |||| |||||||||| ||  |
Db        228 AAGAGCAAAGCCAAGATTGGAAATGACACGATTCTTTTCACATTGTATAAAAAGGAGCCA 287

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |   ||||||| | |||||||||| |||||||||| |||||||||||||| ||||| |||
Db        288 GTTCTGTGGGATAGCCTTTCTGTGCCGGGTGTTGATAAAGAGATGATGCAGAGAATAAGA 347

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              |||||||||| || |||||||| |||| ||||||||| || ||||||||||||||||
Db        348 GAAAAATCTATCTTGCAAGCACAGGAGAAAGCAAAAGAGGCCACAGAAGCAAAAGCTGTT 407

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              || ||||| ||||| || | |||||||||| | |  |||||||||||||||||||||||
Db        408 GCCAAGCGAGAAGACCAGAGATACGCACTAGGCGAGATGATGAAGATTGAAGAAGAAGAG 467

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||| | |||||| |||||||||||||||||| ||| || |||| | ||| |||
Db        468 AGGAAAAAACTCGAAGATCTGAAAGAAAATGAACGGAAAAAGGCAACTAGCGAATTAGAA 527

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              || |||||||||| ||||| |||||||| |  ||||||| | | |||||  | |||||
Db        528 GCGTGGAAAGAATGTCAAAAGAAAGCTGACGGACAAAAAAGAGTCCAGAGGAAGGAGAA- 586

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |  |   | | |||||||  |   |||| |  |||   |||||  |  | || |
Db        587 --ACCGCTCGAGGGAAAGCAAGCT---GAAGAGACCAAAGCTCTAAAACCTCGGGGTTTG 641

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              | | ||     || | |     ||  |  ||| |||||| |||| ||||| ||||||
Db        642 CCCCGGAAGGCCCCACCCACTCGCCTCCCCACAAGAGGGAGGAATTGGGAAAACATATTT 701

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              |||||||||||||||||||||  | ||||| |||||||| | |||||||||| |||||
Db        702 CCTGAGAAGTTAAAGGAAGACAGAGTCCCTGCGCCTCGCTCCGCTGGCAGTATTCAAATC 761

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              | |||||||||||||| |||||||||| ||||| ||||| ||||| ||||||||||||
Db        762 AGCTTTACCCCTCGAGTGTTCCCAACAGCACTTCGGGAATCCCAAGTCGCAGAAGAGGAG 821

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              |||||||| || |||||||| || ||||| ||||| |||| |||||| |  |||| |
Db        822 GAGTGGCTGCATAAACAAGCAGAAGCACGGAGAGCCATGAGCACTGACCTTCCTGAGTTC 881

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              |  || |||||||||||||||| | ||| ||||| |||||||| || || ||||||||||
Db        882 TTTGACTTAAAAGAAGAAGAGAGGAATCCAGACTGGTTGAAAGACAAGGGAAACAAATTG 941

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| |||||||||||||||||  | ||||||||||||||||||| |||| ||
Db        942 TTTGCAACAGAAAACTATTTGGCAGCGGTTGATGCATATAATTTAGCCATACGACTGAAC 1001

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||  ||| |||||||||||||| |||||||||||||||||  ||| ||||||| ||
Db       1002 TGTAAGATCCCATTATTGTATTTGAATCGGGCTGCTTGCCACCTCAAATTAAAAAACCTA 1061

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              |||||||| || || || |||||||||||||||| || |||||||  |||||||| | |||
Db       1062 CACAAGGCCATCGAGGACTCTTCTAAGGCACTAGAGTTATTGACACCACCTGTTGCTGAC 1121

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              |||||  |||||||||||||||||||| || |||||  || |||||  |||||||||||||||
Db       1122 AATGCCAATGCAAGAATGAAGGCACACGTCCGACGAGGGACAGCGTTCTGTCAACTAGAA 1181

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              |||||||| |||||| | || |||||||||| |||||||||||||| ||| |||||| |
Db       1182 TTGTATGTTGAAGGCTTGCAAGATTATGAAGCTGCACTTAAGATTGACCCAGCCAACACA 1241

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              |||||||  |||||  |||||||||||||||| ||||||||| || | ||| || |||
Db       1242 GTTGTACAGAACGATGCAGAGAAGATTCGGAATATAATTCAAGGGACGGCACTGAAGTCT 1301
```

```
RESULT 10
ADC32116
ID   ADC32116 standard; cDNA; 608 BP.
XX
AC   ADC32116;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human novel cDNA contig sequence, SEQ ID NO:2198.
XX
KW   Human; diagnostic; drug screening; forensics; gene mapping;
KW   biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW   neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW   ulcers; osteoporosis; autoimmune disease; cancer;
KW   molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW   neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW   antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW   gene therapy; chromosome 15q21.3; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003029271-A2.
XX
PD   10-APR-2003.
XX
PF   24-SEP-2002; 2002WO-US030474.
XX
PR   24-SEP-2001; 2001US-0324631P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang TY,  Zhang J,  Ren F,  Xue AJ,  Zhao QA,  Wang J,  Wehrman T;
PI   Zhou P,  Ghosh M,  Wang D,  Ma Y,  Asundi V,  Wang Z,  Weng G;
PI   Haley-Vicente D,  Drmanac RT;
XX
DR   WPI; 2003-371981/35.
DR   P-PSDB; ADC32883.
XX
PT   New polynucleotide and polypeptide useful for diagnosing, preventing or
PT   treating conditions such as neurodegenerative diseases, anemias, platelet
PT   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT   cancer.
XX
PS   Example 2; SEQ ID NO 2198; 1185pp; English.
XX
CC   The invention relates to 971 novel human cDNA sequences (ADC29919-
CC   ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC   invention also relates to nucleic acid sequences over 99% identical with
CC   the novel human cDNAs. The invention additionally encompasses expression
CC   vectors and host cells comprising a nucleic acid of the invention; the
CC   recombinant production of a polypeptide of the invention; an antibody
CC   against a polypeptide of the invention; a method of detecting
CC   polynucleotides or polypeptides of the invention; and methods of
CC   identifying a compound which binds to a polypeptide of the invention. The
CC   invention further discloses methods of peventing, treating or
CC   ameliorating a medical condition; kits comprising polynucleotide probes
CC   and/or monoclonal antibodies for carrying out the methods of the
CC   invention; methods for the identification of compounds that modulate the
CC   expression or activity of the polynucleotide and/or polypeptide; and 767
CC   contig sequences corresponding to the cDNA sequences of the invention
CC   (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC   -ADC33394). The nucleic acids and polypeptides of the invention are
CC   useful in diagnostics, drug screening, forensics, gene mapping, in the
CC   identification of mutations responsible for genetic disorders or other
CC   traits, for assessing biodiversity, and in producing many other types of
CC   data and products dependent on DNA and amino acid sequences. They are
CC   also used for treating diseases such as Parkinson's disease, Alzheimer's
CC   disease and other neurodegenerative diseases, anaemia, platelet
CC   disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC   cancer. The nucleic acids may also be used as hybridisation probes or
CC   primers, and in the recombinant production of a protein. The polypeptides
CC   are also useful in generating antibodies, as molecular weight markers,
CC   and as food supplements. The present sequence represents a human contig
CC   sequence used in an example of the invention. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 608 BP; 230 A; 103 C; 135 G; 140 T; 0 U; 0 Other;

  Query Match           41.5%;  Score 524;  DB 10;  Length 608;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-103;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAA 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAA 60

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 120

Qy        644 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 180

Qy        704 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 240

Qy        764 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 300

Qy        824 CTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGG 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGG 360

Qy        884 ATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATT 943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATT 420

Qy        944 TAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACC 1003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACC 480

Qy       1004 TAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAG 1047
              |||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAG 524


RESULT 11
ACH35463
ID   ACH35463 standard; cDNA; 488 BP.
XX
AC   ACH35463;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human endothelial cell cDNA #3596.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
```

```
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 22675; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 488 BP; 141 A; 105 C; 126 G; 116 T; 0 U; 0 Other;

  Query Match           29.3%;  Score 369.8;  DB 9;  Length 488;
  Best Local Similarity   99.5%;  Pred. No. 2.4e-70;
  Matches  371;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         35 AGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGG 94
               ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         28 AGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGAGCGTCAGAGACACGG 87

Qy         95 ACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCAT 154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 ACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCAT 147

Qy        155 TTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTG 214
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTG 207

Qy        215 TCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTG 274
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTG 267

Qy        275 ACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAA 334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAA 327

Qy        335 AAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTG 394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTG 387

Qy        395 TCATGATGAAGAT 407
               |||||||||||| |
Db        388 TCATGATGAAGCT 400


RESULT 12
ACH23091
ID   ACH23091 standard; cDNA; 458 BP.
XX
AC   ACH23091;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human adult ovary cDNA #1471.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US2003073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT,  Labat I,  Stache-Crain B,  Dickson MC,  Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PT   New polynucleotide sequences obtained from various cDNA libraries, useful
PT   as hybridization probes, as oligomers for PCR, for chromosome and gene
PT   mapping, in the recombinant production of protein, or in generating
PT   antisense DNA or RNA.
XX
PS   Claim 1; SEQ ID NO 10303; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 458 BP; 170 A; 86 C; 97 G; 102 T; 0 U; 3 Other;

  Query Match           26.6%;  Score 336;  DB 9;  Length 458;
  Best Local Similarity   91.1%;  Pred. No. 5e-63;
  Matches  368;  Conservative    0;  Mismatches   35;  Indels    1;  Gaps    1;

Qy        639 GAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCG 698
              ||||||  | || |||||||||||||||| |||||||||||||||||| |||| ||||
Db         56 GAGAAAGACTGACAATATATTTACTGAGACGTTAAAGGAAGACAGTATTACTGCACCTCT 115

Qy        699 CTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGA 758
              | ||   ||  ||      | ||||   |||  ||||| ||||| | |||  |||| |||
Db        116 GTGTGCATGCGCTA-CCCCACCAACCCGACCAGTCGAGAATTCCAATCAGTGCTTCCTGA 174

Qy        759 ATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAAT 818
              ||  |||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATTGCAAGTAACACAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAAT 234

Qy        819 GAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTT 878
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTT 294

Qy        879 GAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATA 938
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATA 354

Qy        939 TAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTG 998
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTG 414

Qy        999 CCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTT 1042
                |||||||||||||||||||||||||||||||||||||||||||
Db        415 CCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTT 458


RESULT 13
AAS70018
ID   AAS70018 standard; cDNA; 1383 BP.
XX
AC   AAS70018;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #5822.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG05831.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 5822; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (I) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC   coding sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1383 BP; 399 A; 343 C; 337 G; 304 T; 0 U; 0 Other;

  Query Match           19.4%;  Score 245.2;  DB 5;  Length 1383;
  Best Local Similarity   91.8%;  Pred. No. 3e-43;
  Matches  259;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy        159 TTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTT 218
                || | |  |    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 TTTTACCACGCACGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTT 755

Qy        219 CACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAA 278
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        756 CACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAA 815

Qy        279 AGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGA 338
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        816 AGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGA 875

Qy        339 AGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCAT 398
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        876 AGCTACAGAAGCAAAAGCTGCAGCAAAGCGCGAAGATCAAAAATACGCACTAAGTGTCAT 935

Qy        399 GATGAAGATTGAAGAAGAAGAGGAAAAAAATAGAAGATAT 440
                ||||||||||  |||  | |   ||||||| || |   | ||
Db        936 GATGAAGATTTCAGACAATGTATGGAAAAACATTGTGAAAAT 977


RESULT 14
ADB16927
ID   ADB16927 standard; DNA; 49806 BP.
XX
AC   ADB16927;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Human DYXC1 DNA, chromosomal gene region nucleotides 50001-100000.
XX
KW   ds; human; DYXC1; dyslexia; neurological disorder; chromosome 15q21;
KW   reading disability; phonological processing; rapid naming;
KW   verbal short-term memory.
XX
OS   Homo sapiens.
XX
PN   WO2003068814-A1.
XX
PD   21-AUG-2003.
XX
PF   12-FEB-2003; 2003WO-FI000110.
XX
PR   12-FEB-2002; 2002US-0355782P.
XX
PA   (LICN ) LICENTIA LTD.
XX
PI   Kere J,  Taipale M,  Nopola-Hemmi J,  Kaminen N;
XX
DR   WPI; 2003-646482/61.
XX
PT   New isolated, purified DYXC1 nucleic acid for studying brain processes,
PT   e.g. reading, phonological processing, rapid naming or verbal short-term
PT   memory, or for diagnosing dyslexia or assessing the predisposition to
PT   dyslexia.
XX
PS   Claim 27; Page 69-83; 135pp; English.
XX
CC   This invention relates to a novel isolated human gene DYXC1 that is
CC   functionally related to dyslexia, more particularly it describes single
CC   nucleotide polymorphisms thought to predispose an individual in to
CC   developing dyslexia. This is a neurological disorder with a genetic basis
CC   (DYXC1 has been isolated to chromosome 15q21), which manifests itself as
CC   a specific reading disability. Specifically, DYXC1 is can be useful in
```

```
CC   study of brain processes such as reading, phonological processing, rapid
CC   naming and verbal short-term memory. Accordingly, the present invention
CC   describes methods and materials for analysing allelic variations in the
CC   DYXC1 gene, and also provides DYXC1 as an antigen for the production of
CC   antibodies used in the diagnosis of dyslexia. This polynucleotide is the
CC   partial genomic sequence of the human DYXC1 chromosomal region
CC   (nucleotides 50001-100000) of the invention.
XX
SQ   Sequence 49806 BP; 12975 A; 10105 C; 10577 G; 15977 T; 0 U; 172 Other;

  Query Match            19.0%;  Score 239.4;  DB 9;  Length 49806;
  Best Local Similarity   92.3%;  Pred. No. 1.3e-41;
  Matches  252;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy        398 TGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGA 457
              ||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32438 TGTTATAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGA 32497

Qy        458 TAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAA 517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32498 TAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAA 32557

Qy        518 AAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAA 577
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      32558 AAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAA 32617

Qy        578 AAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAG 637
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32618 AAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAG 32677

Qy        638 GGAGAAATTCAGAAAATATATTTACTGAGAAGT 670
              | | ||     | ||| | || | |    |||
Db      32678 GTAAAATGATTAAGAATTTCTTGAGTAGTGAGT 32710


RESULT 15
AAC26799
ID   AAC26799 standard; cDNA; 313 BP.
XX
AC   AAC26799;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 30874.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-00200610.
XX
PR   26-FEB-1999;   99US-0122487P.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 30874; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC   ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC   to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC   diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors
XX
SQ   Sequence 313 BP; 112 A; 58 C; 60 G; 76 T; 0 U; 7 Other;

  Query Match            14.2%;  Score 178.8;  DB 3;  Length 313;
  Best Local Similarity   94.8%;  Pred. No. 5e-29;
  Matches  201;  Conservative    5;  Mismatches    3;  Indels    3;  Gaps    2;

Qy        844 GATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTT 903
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTT 60

Qy        904 GCAACGGAAAACTATTTGGCAGCTA--TCAATGCATAT-AATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||:|::  :||:|||||| ||||||||||||||||||||||||
Db         61 GCAACGGAAAACTATTTGGCARCWWATCVAAWGCATATAAATTTAGCCATAAGACTAAAT 120

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 180

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACT 1052
              |||||||||||||||||||||||||||||| | |
Db        181 CACAAGGCTATTGAAGATTCTTCTAAGGAATT 212


Search completed: July  8, 2005, 01:54:57
Job time : 759 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:40:13 ; Search time 4597 Seconds
(without alignments)
10457.945 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1034.8	81.9	1600	3	BC017392	BC017392 Homo sapi
2	812.6	64.3	844	6	CD107587	CD107587 AGENCOURT
3	676.4	53.6	933	6	CD358543	CD358543 AGENCOURT
4	667.8	52.9	793	7	CO738958	CO738958 SlLE04c21
5	567.4	44.9	792	4	BG771796	BG771796 602720472
6	553	43.8	788	2	BE564350	BE564350 601343161
7	519	41.1	683	5	BP460416	BP460416 BP460416
8	513.8	40.7	1190	3	AK005832	AK005832 Mus muscu
9	510.4	40.4	1131	5	BQ217312	BQ217312 AGENCOURT
10	508	40.2	917	6	BY705981	BY705981 BY705981
11	488.6	38.7	545	2	BF216970	BF216970 601884034
12	474.4	37.6	535	7	CN478982	CN478982 UI-CF-FN0
13	449	35.6	796	7	CK603982	CK603982 AGENCOURT
14	434.2	34.4	468	6	CB297042	CB297042 12B22058_
15	427	33.8	1183	2	BF207765	BF207765 601861861
16	422	33.4	462	7	CN429788	CN429788 170006000
17	414.4	32.8	417	1	AI024221	AI024221 ov81e02.s
18	409.6	32.4	988	4	BI517373	BI517373 603041624
19	408	32.3	559	5	BU607404	BU607404 UI-CF-FN0
20	405	32.1	823	5	BU567908	BU567908 AGENCOURT
21	395.6	31.3	1153	3	AK015345	AK015345 Mus muscu
22	389.2	30.8	701	4	BG482634	BG482634 602502783
23	385.8	30.5	651	5	BU852580	BU852580 AGENCOURT
24	376.8	29.8	844	7	CK598370	CK598370 AGENCOURT
c 25	374.4	29.6	772	6	CA422987	CA422987 UI-H-FL0-
26	371.8	29.4	665	4	BG540324	BG540324 602568825
27	364	28.8	573	2	BE972748	BE972748 601652170
28	363.2	28.8	775	4	BG206612	BG206612 RST26063
29	350.8	27.8	477	6	CB545994	CB545994 AMGNNUC:M
30	344.4	27.3	744	4	BG192162	BG192162 RST11269
31	341	27.0	752	4	BG183140	BG183140 RST2158 A
c 32	325	25.7	325	1	AI783611	AI783611 tz99c01.x
33	308.4	24.4	375	2	AW481240	AW481240 35107 MAR
34	301.8	23.9	687	4	BG242087	BG242087 602354631
35	298	23.6	764	5	BU307742	BU307742 603539279
36	296.2	23.5	767	5	BX867086	BX867086 BX867086
37	295	23.4	855	2	BF248143	BF248143 601859338
c 38	293.2	23.2	727	4	BJ056709	BJ056709 BJ056709
39	291.8	23.1	478	2	BB615485	BB615485 BB615485
40	291.8	23.1	478	3	AK076862	AK076862 Mus muscu
41	288.4	22.8	543	7	CK137348	CK137348 mak06b02.
42	283.8	22.5	470	1	AV266346	AV266346 AV266346
43	273.2	21.6	375	2	AW415915	AW415915 50508 MAR
44	270.8	21.4	617	5	BU267200	BU267200 603819026
45	269.4	21.3	370	1	AA967396	AA967396 ua05c07.r

ALIGNMENTS

RESULT 1
BC017392
LOCUS BC017392 1600 bp mRNA linear HTC 23-AUG-2002
DEFINITION Homo sapiens, Similar to RIKEN cDNA 1700010I24 gene, clone IMAGE:4081622, mRNA.
ACCESSION BC017392
VERSION BC017392.1 GI:19263480
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 32 Row: k Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction
This clone has the following problem: frame shifted.
FEATURES Location/Qualifiers
source 1. .1600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4081622"
/tissue_type="Bladder, carcinoma"
/clone_lib="NIH_MGC_53"
/lab_host="DH10B"

```
                     /note="Vector: pDNR-LIB"
ORIGIN

  Query Match            81.9%;  Score 1034.8;  DB 3;  Length 1600;
  Best Local Similarity  99.7%;  Pred. No. 9.3e-215;
  Matches 1047;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         53 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGTAGACTGCGGTCTTTCTGTCT 112

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 172

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 232

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 292

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 352

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 412

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 472

Qy        421 AGG-AAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGA 479
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGA 532

Qy        480 AGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAA 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAA 592

Qy        540 ATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCT 599
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        593 ATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCT 652

Qy        600 TACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATT 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATT 712

Qy        660 TACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAAT 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 TACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAAT 772

Qy        720 CAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGA 779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 CAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGA 832

Qy        780 GGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACT 839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACT 892

Qy        840 TTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATT 899
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 TTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATT 952

Qy        900 GTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAA 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAA 1012

Qy        960 TAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTT 1019
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1013 TAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTT 1072

Qy       1020 ACACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1073 ACACAAGGCTATTGAAGATTCTTCTAAGGC 1102


RESULT 2
CD107587
LOCUS       CD107587                844 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  AGENCOURT_14021179 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30368368 5', mRNA sequence.
ACCESSION   CD107587
VERSION     CD107587.1  GI:30760665
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 844)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM431  row: a  column: 17
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1. .844
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30368368"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match            64.3%;  Score 812.6;  DB 6;  Length 844;
  Best Local Similarity  99.3%;  Pred. No. 2.2e-166;
  Matches  837;  Conservative    0;  Mismatches    4;  Indels    2;  Gaps    2;

Qy         18 CGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGT 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGT 60

Qy         78 GTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCC 120

Qy        138 ATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGAT 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGAT 180

Qy        198 TGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCT 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCT 240

Qy        258 TTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 TTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACA 300

Qy        318 AGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCA 360

Qy        378 AAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGA 420

Qy        438 TATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCA 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCA 480

Qy        498 AAGAAAAGCTGAGGAGC-AAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAA 556
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAA 540

Qy        557 AGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCAT 616
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCAT 600

Qy        617 CTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGG 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGG 660

Qy        677 AAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAG 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAG 720

Qy        737 TATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAAC 796
              |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||
Db        721 TATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAAAAGAAGAGGAGTGGCTACACAAAC 780

Qy        797 AAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTT-GCGATTTAAAAGAA 855
              |||||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||
Db        781 AAGCTGAGGCACGAAGAGCCATGAATACTGACATAGCTGAACTTTGGCGATTTAAAAGAA 840

Qy        856 GAA 858
              |||
Db        841 GAA 843
```

RESULT 3
CD358543
LOCUS CD358543 933 bp mRNA linear EST 29-MAY-2003
DEFINITION AGENCOURT_14255831 NIH_MGC_180 Homo sapiens cDNA clone IMAGE:30386203 5', mRNA sequence.
ACCESSION CD358543
VERSION CD358543.1 GI:31129978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM454 row: h column: 20
High quality sequence start: 303
High quality sequence stop: 666.

FEATURES Location/Qualifiers
source 1. .933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30386203"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 53.6%; Score 676.4; DB 6; Length 933;
Best Local Similarity 96.5%; Pred. No. 1e-136;
Matches 723; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

```
Qy        159 TTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTT 218
              |||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||
Db         33 TTATGCTCCCATAGACTACCAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTT 92

Qy        219 CACCTTGTATAAAAAA-GAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACA 277
              |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db         93 CACCTTGTATAAAAAACCAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACA 152

Qy        278 AAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAG 337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 AAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAG 212

Qy        338 AAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCA 397
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 AAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCA 272

Qy        398 TGATGAAGATTGAAGAAGAAGAGAGGAAAAAAA-TAGAAGATATGAAAGAAAATGAACGG 456
              |||||||||||||||||||||||||||||||| |     ||||||||||||||||||||
Db        273 TGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATTTTTTTATATGAAAGAAAATGAACGG 332

Qy        457 ATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAA 516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 ATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAA 392

Qy        517 AAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGA 576
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        393 AAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGA 452

Qy        577 AAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAA 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 AAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAA 512

Qy        637 GGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCT 696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCT 572

Qy        697 CGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGT 756
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 CGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGT 632

Qy        757 GAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCA 816
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 GAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCA 692

Qy        817 ATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCC--AGAAT 874
              |||||||||||||||||||||||||||||||||||||||||||||| || ||||    |
Db        693 ATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAANAGGACCCCAGAATG 752

Qy        875 GGTTGAAGGATAAAGGAAACAAATTGTTT 903
              ||||||||||||||| ||  || | | |
Db        753 GGTTGAAGGATAAAGGGAAACAAATTTGT 781
```

```
RESULT 4
CO738958
LOCUS       CO738958                 793 bp    mRNA    linear   EST 29-JUL-2004
DEFINITION  SlLE04c21p08f1 squirrel embryo library 1 Spermophilus lateralis
            cDNA clone 21p08 5', mRNA sequence.
ACCESSION   CO738958
VERSION     CO738958.1  GI:50826228
KEYWORDS    EST.
SOURCE      Spermophilus lateralis (golden-mantled ground squirrel)
  ORGANISM  Spermophilus lateralis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
            Spermophilus.
REFERENCE   1  (bases 1 to 793)
  AUTHORS   Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
            Rogers,J. and Cossins,A.R.
  TITLE     Microarray analysis of transcriptional changes during hibernation
            in the golden mantled ground squirrel, Spermophilus lateralis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Andrew R. Cossins
            Laboratory for Environmental Gene Regulation
            University of Liverpool
            School of Biological Sciences, The Biosciences Building, Crown
            Street, Liverpool, United Kingdom, L69 7ZB
            Tel: +44(0)151-795-4510
            Fax: +44(0)151-795-4431
            Email: cossins@liv.ac.uk
            Vector has been trimmed from this EST.
            Plate: 21  row: p  column: 08
            Seq primer: pflc T7 (5'-AATACGACTCACTATAGGG-3')
            High quality sequence stop: 793.
FEATURES             Location/Qualifiers
     source          1. .793
                     /organism="Spermophilus lateralis"
                     /mol_type="mRNA"
                     /db_xref="taxon:76772"
                     /clone="21p08"
                     /sex="male and female"
                     /tissue_type="embryo"
                     /dev_stage="embryonic"
                     /lab_host="E.coli Electromax DH10B"
                     /clone_lib="squirrel embryo library 1"
                     /note="Vector: pFLC; Site_1: SalI GTCGAG; Site_2: BamHI
                     GGATCC; Normalized and subtracted cDNA library prepared
                     from embryos"
ORIGIN

  Query Match           52.9%;  Score 667.8;  DB 7;  Length 793;
  Best Local Similarity   90.3%;  Pred. No. 7.5e-135;
  Matches  714;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy        346 GAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAG 405
              ||||||||||| |||||||||| ||||||||||||||  |||||| |||||||||| ||
Db          2 GAAGCAAAAGCTACAGCAAAGCGAGAAGATCAAAAATATACACTAAATGTCATGATGCAG 61

Qy        406 ATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCC 465
              |||||||||||||||||||| || ||||||||||||||||||||||||| ||||||||
Db         62 ATTGAAGAAGAAGAGAGGAAGAAGATAGAAGATATGAAAGAAAATGAACGAATAAAAGCA 121

Qy        466 ACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATT 525
              ||||||| |||||||||||||||||| | |||||||||||||| || | ||||| ||||
Db        122 ACTAAAGAATTGGAAGCCTGGAAAGAGTGTCAAAGAAAAGCTGAAGAACGAAAAAGAATT 181

Qy        526 CAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATA 585
              |||| |||||||||||||| ||||| |||||||||||||| | |||| |||||||||| ||
Db        182 CAGAAAGAAGAGAAATTACGTCAACAAGAAAAGCAAATTGACGAAGAAAGAAAAAAATTA 241

Qy        586 AAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAAT 645
              | | |||| | ||| | || |||||   |||||| |||||||||  ||||||||||||||
Db        242 AGACATAAAAATCTCATTAAAAATTCAACATCTAAAAATCTTGCCACAAAAGGGAGAAAT 301

Qy        646 TCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTT 705
              |||||||||||||| | | |||||||||||||||| |||||||||  ||||||||||||
Db        302 TCAGAAAATATATTTCGGGGAAGTTAAAGGAAGACTGTATTCCTGCACCTCGCTCTGTT 361

Qy        706 GGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAA 765
              || | |||||||||||||| |||||||||||||||||||| || || ||||||||| ||
Db        362 GGAAATATTAAAATCAACTTCACCCCTCGAGTATTCCCAACTGCACTCCGTGAATCAAAA 421

Qy        766 GTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACT 825
              |||||||||| |||||||||||||| |||||||||||||||||||| ||||||||||||
Db        422 GTAGCAGAAGAAGAGGAGTGGCTACATAAACAAGCTGAGGCACGAAAAGCAATGAATACT 481

Qy        826 GACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGAT 885
              || ||  ||||| ||||||| |||||||||| ||||||||||||||||||||||||||||
Db        482 GATATTCCTGAATTTGCGACTTAAAAGAAGACGAAAAGAACCCAGAATGGTTGAAGGAG 541

Qy        886 AAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTA 945
              |||||||||||||||||||||||| ||  |||||||||||||||| |||||||| ||||||
Db        542 AAAGGAAACAAATTGTTTGCAACAGAGGACTATTTGGCAGCTATTAATGCATACAATTTA 601

Qy        946 GCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTA 1005
              |||||||||| ||||| ||||| || ||||| |||||||| ||||||||||||||||||
Db        602 GCCATAAGACTGAATAACAAGATTCCTCTATTATATTTGAATCGGGCTGCTTGCCACCTA 661

Qy       1006 AAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATG 1065
              |||||||||||||||||||||| |||||||||||||||||| ||||||||| |||
Db        662 AAACTAAAAAACTTACACAAGGCCATTGAAGATTCTTCTAAGGCTCTGGAATTACTGACA 721

Qy       1066 CCACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCA 1125
              |||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 CCACATGTTGCAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCA 781

Qy       1126 TTCTGTCAACT 1136
              ||||||||||
Db        782 TTCTGTCAACT 792


RESULT 5
BG771796
LOCUS       BG771796                 792 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602720472F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837505 5',
            mRNA sequence.
ACCESSION   BG771796
VERSION     BG771796.1  GI:14082449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10769  row: p  column: 18
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4837505"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
```

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 44.9%; Score 567.4; DB 4; Length 792;
Best Local Similarity 91.5%; Pred. No. 5.5e-113;
Matches 671; Conservative 0; Mismatches 46; Indels 16; Gaps 6;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 96

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 156

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 216

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        217 AGCAGCAAAGCAAAGATTGGGAATGACACCATGGTCTTCACCTTGTATAAAAAAGAAGCG 276

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 336

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 396

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 456

Qy        421 AGGAAAAAA-ATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGA 479
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 AGGAAAAAACATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGA 516

Qy        480 AGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA--TTCAGAGAGAAGAG 537
              |||||||||||||||||||||||||||||||||||||||||||  |      ||||||||
Db        517 AGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGGAGAAGAG 576

Qy        538 AAATTATGTCAAAAAGA------AAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATAT 591
              |||||||||||||||||      |||  ||   ||||| ||| | |  ||   ||||||
Db        577 AAATTATGTCAAAAAGAAAAAGCAAAATTAAACGAAGAAAGAACATACCAATACAAATAT 636

Qy        592 AAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTG---CTCCAAAAGGGAGAAATTCA 648
              |||||||||||||||| ||||||||||||||||||    | |  || ||||||||||||
Db        637 AAGAGTCTTACTAGAAACTTGGCATCTAGAAATCTTGGCTCCCAGAAGGGGAGAAATTCC 696

Qy        649 G--AAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCC--TGCTCCTCGCTCTGT 704
              |  ||      ||  |||||||| || | |  ||||||||||      | |||||||||
Db        697 GCCCAATATATTTACCTGAGAAGTAAAGGAAGAACAGTATTCCCTGCTCCGTCGCTCTGT 756

Qy        705 TGGCAGTATTAAA 717
              ||||||  |||||
Db        757 TGGCAGACTTAAA 769
```

RESULT 6
BE564350
LOCUS BE564350 788 bp mRNA linear EST 15-AUG-2000
DEFINITION 601343161F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3685335 5', mRNA sequence.
ACCESSION BE564350
VERSION BE564350.1 GI:9808070
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 788)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM373 row: i column: 16
High quality sequence stop: 556.
FEATURES Location/Qualifiers
source 1. .788
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3685335"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_53"
/note="Organ: bladder; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN

Query Match 43.8%; Score 553; DB 2; Length 788;
Best Local Similarity 94.6%; Pred. No. 7.6e-110;
Matches 594; Conservative 0; Mismatches 31; Indels 3; Gaps 2;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 82

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 142

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 202

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 262

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 322

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 382

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||||
Db        383 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGACGACGAAGAG 442

Qy        421 AGG--AAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGG 478
              |||  ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 AGGACAAACAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGG 502

Qy        479 AAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGA 538
              ||||||||||||||||||||||||||||||||||  |||||||| || ||||||||||||
Db        503 AAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGNCAAAAAAAAATCCGAGAGAAGAGA 562

Qy        539 AATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTC 598
              | |  ||||  | |||||||||  | |   | |     || |  | ||||||||||||||
Db        563 ACTCTTGTC-GACAGAAAAGCACTTCACGAAGGCCGACAACACCAATAAATATAAGAGTC 621

Qy        599 TTACTAGAAATTTGGCATCTAGAAATCT 626
              |||||||||| |||| ||||||||||||
Db        622 TTACTAGAAATGTGGCTTCTAGAAATCT 649


RESULT 7
BP460416
LOCUS       BP460416                683 bp    mRNA    linear   EST 31-DEC-2003
DEFINITION  BP460416 full-length enriched swine cDNA library, adult ovary Sus
            scrofa cDNA clone OVRM10147G01 5', mRNA sequence.
ACCESSION   BP460416
VERSION     BP460416.1  GI:40476478
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 683)
  AUTHORS   Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
            Okumura,N., Hamasima,N. and Awata,T.
  TITLE     PEDE (Pig EST Data Explorer): construction of a database for ESTs
            derived from porcine full-length cDNA libraries
  JOURNAL   Nucleic Acids Res. 32 (1), D484-D488 (2004)
COMMENT     Contact: Hirohide Uenishi
            Animal Genome Laboratory, Genome Research Department
            National Institute of Agrobiological Sciences
            2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
            Tel: +81-29-838-8627
            Fax: +81-29-838-8627
            Email: huenishi@affrc.go.jp
            EST project with full-length enriched cDNA libraries carried out in
            Animal Genome Research Program (Japan) by National Institute of
            Agrobiological Sciences and STAFF-Institute
            Single pass sequencing of clones derived from oligo-capped cDNA
            library
            Vector sequences were eliminated by RepeatMasker version 2002/07/13
            and crossmatch version 0.990319
            Low quality bases were trimmed based on the quality values.
FEATURES             Location/Qualifiers
     source          1. .683
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /clone="OVRM10147G01"
                     /tissue_type="ovary"
                     /dev_stage="adult"
                     /clone_lib="full-length enriched swine cDNA library, adult
                     ovary"
ORIGIN

  Query Match             41.1%;  Score 519;  DB 5;  Length 683;
  Best Local Similarity   90.2%;  Pred. No. 1.9e-102;
  Matches  555;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy        539 AATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTC 598
              |||||  |||| ||| |||||||||| ||||||  |||||||| ||||   ||| || | |
Db          1 AATTACATCAACAAGGAAAGCAAATTGAAGAAGAGAGAAAAAAAATTAAAGCGTAAAAGCC 60

Qy        599 TTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATAT 658
              ||||||||||| |||||||| ||||||   |||| ||||||||||||||||||||||||
Db         61 TTACTAGAAATTCAGCATCTAGATATCTTGACACAAATGGGAGAAATTCAGAAAATATAT 120

Qy        659 TTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAA 718
              ||  |||||||||||||||||||||| |||||||||||||||||||||| ||||||||
Db        121 TTTTTGAGAAGTTAAAGGAAGACAGTATCCCTGCTCCTCGCTCTGTTGGCAATATTAAAA 180

Qy        719 TCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGG 778
              |||||||||||||||||||||||||| || || || |||||||||||||| ||||| |
Db        181 TCAACTTTACCCCTCGAGTATTCCCAACCGCCCTCCGGGAATCACAAGTAGCGGAAGAAG 240

Qy        779 AGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAAC 838
              |||||||||||||||||||||||| |||||||||||||| |||||||||||||| ||  ||||
Db        241 AGGAGTGGCTACACAAACAAGCAGAGGCACGAAGGGCAATGAATACTGAAATTCCTGAGT 300

Qy        839 TTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAAT 898
              || | |||||||||||||| |||||||||||||||||||||||| |||||||||||| |
Db        301 TTAGTGATTTAAAAGAAGAGGAAAAGAACCCAGAATGGTTGAAGGACAAAGGAAACAAGT 360

Qy        899 TGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAA 958
              |||||||||| |||||||||||||||||||| |||||||| || ||||||||||||||||
Db        361 TGTTTGCAACAGAAAACTATTTGGCAGCTATTAATGCATACAACTTAGCCATAAGACTAA 420

Qy        959 ATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACT 1018
              |||||||||| ||||| |||||||||| | |||||||||||||||||||||||| ||||||
Db        421 ATAATAAGATTCCACTGTTGTATTTGGATCGGGCTGCTTGCCACCTAAAACTGAAAAACC 480

Qy       1019 TACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAG 1078
              ||||||| || |||||||||||||||||||| | |||||||| |  |||||||||| |||
Db        481 TACACAAAGCCATTGAAGATTCTTCTAAGGCATTAGAATTATTAACACCACCTGTTGCAG 540

Qy       1079 ACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAG 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAG 600

Qy       1139 AATTGTATGTAGAAG 1153
              |||||||||||||||
Db        601 AATTGTATGTAGAAG 615


RESULT 8
AK005832
LOCUS       AK005832                1190 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus adult male testis cDNA, RIKEN full-length enriched
            library, clone:1700010I24 product:EKN1 homolog [Homo sapiens], full
            insert sequence.
ACCESSION   AK005832
VERSION     AK005832.1  GI:12838612
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
```

```
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1190)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
            Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
            Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
            Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
            Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
            Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
            Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.jp/) for further
            details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence[5'
            GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA
            was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
            end: SstI. Host: SOLR.
FEATURES             Location/Qualifiers
     source          1. .1190
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:1700010I24"
                     /db_xref="taxon:10090"
                     /clone="1700010I24"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             351. .803
                     /note="unnamed protein product; EKN1 homolog [Homo
                     sapiens] (SPTR|AAL73230, evidence: FASTY, 89.1%ID,
                     41.4%length, match=522)
                     putative"
                     /codon_start=1
                     /protein_id="BAB24264.2"
                     /db_xref="GI:26348135"
                     /translation="MSTDLPEFFDLKEEERNPDWLKDKGNKLFATENYLAAVDAYNLA
                     IRLNCKIPLLYLNRAACHLKLKNLHKAIEDSSKALELLTPPVADNANARMKAHVRRGT
                     AFCQLELYVEGLQDYEAALKIDPANTVVQNDAEKIRNIIQGTALKSRD"
     polyA_signal    1169. .1174
                     /note="putative"
     polyA_site      1190
                     /note="putative"
ORIGIN

  Query Match             40.7%;  Score 513.8;  DB 3;  Length 1190;
  Best Local Similarity   79.5%;  Pred. No. 2.8e-101;
  Matches  634;  Conservative    0;  Mismatches  157;  Indels    6;  Gaps    2;

Qy        464 CCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAA 523
              |  ||||  | ||| ||||| |||||||||| |||||  ||||||||  |  ||||||| |
Db          4 CAACTAGCGAATTAGAAGCGTGGAAAGAATGTCAAAAGAAAGCTGACGGACAAAAAAGAG 63

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAA 583
              | |||||  | |||||   |  |   | | |||||||  | ||||    |  |||
Db         64 TCCAGAGGAAGGAGAA---ACCGCTCGAGGGAAAGCAAGCTGAAGAG---ACCAAAGCTC 117

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 643
              |||||  |  | ||  |  |  | ||     || | |     ||  |  ||| |||||| |
Db        118 TAAAACCTCGGGGTTTGCCCCGGAAGGCCCCACCCACTCGCCTCCCCACAAGAGGGAGGA 177

Qy        644 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 703
              |||  ||||| ||||||  |||||||||||||||||||||||  | |||||  ||||||| |
Db        178 ATTGGGAAAACATATTTCCTGAGAAGTTAAAGGAAGACAGAGTCCCTGCGCCTCGCTCCG 237

Qy        704 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 763
              |||||||||| |||||| |||||||||||||| |||||||||| ||||| ||||| |
Db        238 CTGGCAGTATTCAAATCAGCTTTACCCCTCGAGTGTTCCCAACAGCACTTCGGGAATCCC 297

Qy        764 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 823
              |||| |||||||||||||||||||| || |||||||| || ||||| ||||| ||||  |
Db        298 AAGTCGCAGAAGAGGAGGAGTGGCTGCATAAACAAGCAGAAGCACGGAGAGCCATGAGCA 357

Qy        824 CTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGG 883
              |||||  |  ||||  | |  || |||||||||||||| | ||| ||||| ||||||| |
Db        358 CTGACCTTCCTGAGTTCTTTGACTTAAAAGAAGAAGAGAGGAATCCAGACTGGTTGAAAG 417

Qy        884 ATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATT 943
              | || |||||||||||||||||||| ||||||||||||||||  |  ||||||||||||
Db        418 ACAAGGGAAACAAATTGTTTGCAACAGAAAACTATTTGGCAGCGGTTGATGCATATAATT 477

Qy        944 TAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACC 1003
              |||||||| |||| ||   |||||| ||| |||||||||||||| ||||||||||||||||
Db        478 TAGCCATACGACTGAACTGTAAGATCCCATTATTGTATTTGAATCGGGCTGCTTGCCACC 537

Qy       1004 TAAAACTAAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGA 1063
              | ||| |||||||| |||||||||| || || || ||||||||||||| || |||||||
Db        538 TCAAATTAAAAAACCTACACAAGGCCATCGAGGACTCTTCTAAGGCACTAGAGTTATTGA 597

Qy       1064 TGCCACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAG 1123
                |||||||| | |||||||| |||||||||||||||||||| || ||||| || ||||
Db        598 CACCACCTGTTGCTGACAATGCCAATGCAAGAATGAAGGCACACGTCCGACGAGGGACAG 657

Qy       1124 CATTCTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGA 1183
              | |||||||||||||||||||||||| |||||| | || |||||||||| ||||||||||
Db        658 CGTTCTGTCAACTAGAATTGTATGTTGAAGGCTTGCAAGATTATGAAGCTGCACTTAAGA 717

Qy       1184 TTGATCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAG 1243
              |||| ||| |||||| | ||||||| |  ||||| |||||||||||||| |||||||||
Db        718 TTGACCCAGCCAACACAGTTGTACAGAACGATGCAGAGAAGATTCGGAATATAATTCAAG 777

Qy       1244 GAACAGAACTAAAATCT 1260
              | || | ||| || |||
Db        778 GGACGGCACTGAAGTCT 794


RESULT 9
BQ217312
```

```
LOCUS       BQ217312                1131 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7558431 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6047052
            5', mRNA sequence.
ACCESSION   BQ217312
VERSION     BQ217312.1  GI:20398712
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1131)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13293  row: n  column: 13
            High quality sequence stop: 388.
FEATURES             Location/Qualifiers
     source          1. .1131
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6047052"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             40.4%;  Score 510.4;  DB 5;  Length 1131;
  Best Local Similarity   95.9%;  Pred. No. 1.5e-100;
  Matches  535;  Conservative    0;  Mismatches   21;  Indels    2;  Gaps    1;

Qy         42 GACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTT 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTT 60

Qy        102 CTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTA 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTA 120

Qy        162 TGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCAC 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCAC 180

Qy        222 CTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGA 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGA 240

Qy        282 GATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGC 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGC 300

Qy        342 TACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGAT 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGAT 360

Qy        402 GAAGATTGAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAA 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAA 420

Qy        462 AGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGC--AAAAA 519
              |||||||||||||||||||||||| |||||||||||||||||||||||||  |||||
Db        421 AGCCACTAAAGCATTGGAAGCCTGGACAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAA 480

Qy        520 AAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAA 579
              |||||||| |||||||||||||||||||||||||| ||| |  | | ||  |    | |
Db        481 AAAATTCAAAGAGAAGAGAAATTATGTCAAAAAGGAAAACCCAATTAAAGAAATGAAGGA 540

Qy        580 AAAATAAAATATAAGAGT 597
              |||| ||| || || |||
Db        541 AAAAAAAATTAAAATAGT 558


RESULT 10
BY705981
LOCUS       BY705981                917 bp    mRNA    linear   EST 16-DEC-2002
DEFINITION  BY705981 RIKEN full-length enriched, adult male testis Mus musculus
            cDNA clone 1700010I24 5', mRNA sequence.
ACCESSION   BY705981
VERSION     BY705981.1  GI:27117128
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
            Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
            Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
            Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
            Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,
            Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
            Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
            Direct Submission
```

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .917
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="1700010I24"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male testis"
/note="Site_1: XhoI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. "

ORIGIN

Query Match 40.2%; Score 508; DB 6; Length 917;
Best Local Similarity 79.1%; Pred. No. 4.9e-100;
Matches 630; Conservative 0; Mismatches 160; Indels 6; Gaps 2;

```
Qy        466 ACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATT 525
              ||||  | ||| ||||| |||||||||| |||||  |||||||| |  ||||||| | |
Db          2 ACTAGCGAATTAGAAGCGTGGAAAGAATGTCAAAAGAAAGCTGACGGACAAAAAAGAGTC 61

Qy        526 CAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATA 585
              |||||  | |||||   |   |   | | ||||||| | ||||   |  |||    ||
Db         62 CAGAGGAAGGAGAA---ACCGCTCGAGGGAAAGCAAGCTGAAGAG---ACCAAAGCTCTA 115

Qy        586 AAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAAT 645
              |||  |  | ||  |  |  | ||     || | |     ||  |  ||| |||||| |||
Db        116 AAACCTCGGGGTTTGCCCCGGAAGGCCCCACCCACTCGCCTCCCCACAAGAGGGAGGAAT 175

Qy        646 TCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTT 705
              |  ||||| |||||| |||||||||||||||||||  | |||||  |||||||| | |
Db        176 TGGGAAAACATATTTCCTGAGAAGTTAAAGGAAGACAGAGTCCCTGCGCCTCGCTCCGCT 235

Qy        706 GGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAA 765
              |||||||| |||||| |||||||||||||| |||||||||| ||||| ||||| |||
Db        236 GGCAGTATTCAAATCAGCTTTACCCCTCGAGTGTTCCCAACAGCACTTCGGGAATCCCAA 295

Qy        766 GTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACT 825
              || |||||||||||||||||| || |||||||| || ||||| ||||| ||||  |||
Db        296 GTCGCAGAAGAGGAGGAGTGGCTGCATAAACAAGCAGAAGCACGGAGAGCCATGAGCACT 355

Qy        826 GACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGAT 885
              |||  |  ||||  | |  || |||||||||||||| | |||  |||||  |||||||| ||
Db        356 GACCTTCCTGAGTTCTTTGACTTAAAAGAAGAAGAGAGGAATCCAGACTGGTTGAAAGAC 415

Qy        886 AAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTA 945
              ||  |||||||||||||||||||| ||||||||||||||||  |  ||||||||||||||
Db        416 AAGGGAAACAAATTGTTTGCAACAGAAAACTATTTGGCAGCGGTTGATGCATATAATTTA 475

Qy        946 GCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTA 1005
              ||||||  |||| ||    |||||| ||| |||||||||||| ||||||||||||||||
Db        476 GCCATACGACTGAACTGTAAGATCCCATTATTGTATTTGAATCGGGCTGCTTGCCACCTC 535

Qy       1006 AAACTAAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATG 1065
              ||| |||||||| |||||||||| || || || |||||||||||||| || |||||||
Db        536 AAATTAAAAAACCTACACAAGGCCATCGAGGACTCTTCTAAGGCACTAGAGTTATTGACA 595

Qy       1066 CCACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCA 1125
              ||||||||| | |||||||| |||||||||||||||||||| || ||||| || |||||
Db        596 CCACCTGTTGCTGACAATGCCAATGCAAGAATGAAGGCACACGTCCGACGAGGGACAGCG 655

Qy       1126 TTCTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATT 1185
              |||||||||||| |||||||||| |||||| | || |||||||| | ||||||||||||
Db        656 TTCTGTCAACTAAAATTGTATGTTGAAGGCTTGCAAGATTATGAAACTGCACTTAAGATT 715

Qy       1186 GATCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGA 1245
              || ||| |||||| || |||||| |  ||||| |||||||||| || | ||||||||||
Db        716 GACCCAGCCAACACAAGTGTACAGAACGATGCAGAGAAGATTCCGATTATAATTCAAGGG 775

Qy       1246 ACAGAACTAAAATCTT 1261
              || | ||| || ||||
Db        776 ACGGCACTGAAGTCTT 791
```

RESULT 11
BF216970
LOCUS BF216970 545 bp mRNA linear EST 06-NOV-2000
DEFINITION 601884034F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4102592 5', mRNA sequence.
ACCESSION BF216970
VERSION BF216970.1 GI:11110556
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM976 row: c column: 09
High quality sequence stop: 539.

FEATURES Location/Qualifiers
source 1. .545
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4102592"
/tissue_type="glioblastoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_57"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

```
  Query Match             38.7%;  Score 488.6;  DB 2;  Length 545;
  Best Local Similarity   97.0%;  Pred. No. 7.8e-96;
  Matches  519;  Conservative    0;  Mismatches   14;  Indels    2;  Gaps    2;

Qy        293 GAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAA 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAA 60

Qy        353 AAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAG 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAG 120

Qy        413 AAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAG 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAG 180

Qy        473 CATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAG 532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAG 240

Qy        533 AAGAGAAATTATGTCAAAAAGAAAAGCAAATTAA-AGAAGGAAGAAAAAAAATAAAATAT 591
              |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db        241 AAGAGAAATTATGTCAAAAAGAAAAGCAAATTAACAGAAGAAAGAAAAAAAATAAAATAT 300

Qy        592 AAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAA 651
              |||||||||||||||||| |||||||||||||| ||||||||||||||||||||||||||
Db        301 AAGAGTCTTACTAGAAATTGGGCATCTAGAAATCTGGCTCCAAAAGGGAGAAATTCAGAA 360

Qy        652 AATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGT 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGGCAG 420

Qy        712 ATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCA 771
              | |||   | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATTAAAATCAATTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCA 480

Qy        772 GAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTG 826
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        481 GAAGAGGAGGAGTGG-TACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTG 534
```

RESULT 12
CN478982
LOCUS CN478982 535 bp mRNA linear EST 23-APR-2004
DEFINITION UI-CF-FN0-afu-l-16-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone UI-CF-FN0-afu-l-16-0-UI 3', mRNA sequence.
ACCESSION CN478982
VERSION CN478982.1 GI:46547981
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/cf.html
The following repetitive elements were found in this cDNA sequence: 133-243, >(GAAAA)n#Simple_repeat 493-514, >AT_rich#Low_complexity (matched compliment)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FN0-afu-l-16-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FN0"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-FN0 is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-FN0
TAG_SEQ=GGCTGTAGGC"

ORIGIN

```
  Query Match             37.6%;  Score 474.4;  DB 7;  Length 535;
  Best Local Similarity   98.8%;  Pred. No. 9.6e-93;
  Matches  478;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy        386 CACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAG 445
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 CACTTAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAG 93

Qy        446 AAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAG 505
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 AAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAG 153

Qy        506 CTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTA 565
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 CTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTA 213

Qy        566 AAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATC 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 AAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATC 273

Qy        626 TTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTA 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTA 333

Qy        686 TTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAA 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 TTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAA 393

Qy        746 CAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGG 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 CAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGG 453

Qy        806 CACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGA 865
              ||||||||||||||||||||||||||||||||||||||||||||| || || |||| |
Db        454 CACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAAAAAAAAAAAAAA 513
```

```
Qy        866 ACCC 869
              | |||
Db        514 AGCC 517


RESULT 13
CK603982
LOCUS       CK603982                796 bp    mRNA    linear   EST 22-JAN-2004
DEFINITION  AGENCOURT_17890681 NIH_MGC_238 Rattus norvegicus cDNA clone
            IMAGE:7134374 5', mRNA sequence.
ACCESSION   CK603982
VERSION     CK603982.1  GI:41117319
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 796)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Howard Jacobs
             cDNA Library Preparation: Express Genomics
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15039  row: o  column: 12
            High quality sequence stop: 622.
FEATURES             Location/Qualifiers
     source          1. .796
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="IMAGE:7134374"
                     /tissue_type="testis, pooled"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_238"
                     /note="Organ: testis; Vector: pExpress-1; Site_1: EcoRV;
                     Site_2: NotI; RNA obtained from testis tissue of 8 wk old
                     animal. Tissues were snap-frozen and kept at -80C before
                     RNA extraction and purification (Tri-reagent method). cDNA
                     was primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
                     resulted in an average insert size of 1.9 kb. This primary
                     library is normalized (non-normalized primary library is
                     NIH_MGC_237) and was constructed by Express Genomics
                     (Frederick, MD)"
ORIGIN

  Query Match             35.6%;  Score 449;  DB 7;  Length 796;
  Best Local Similarity   76.9%;  Pred. No. 3.4e-87;
  Matches  574;  Conservative    0;  Mismatches  166;  Indels    6;  Gaps    2;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              |||||  | |  || |||||   |||||||||||||||   | | ||  |||| |||||
Db          4 ATGCCGGTGCGAGTGAGCGAGCTCAGCTGGCAGCAGACACCGGCCGCACTCTTCCTGTCG 63

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||| ||    ||||| ||||| | ||  | || ||||||||   |||||  || |||
Db         64 CTGCCTCTACGGGGCGTCTGCGTGCGCGATGCTGATGTGTTCTGTGGGGAAAGTTACCTG 123

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||| ||||||||||||||||||||||||  ||||| |||||||||||||| |||| |
Db        124 AAGGTTAACTTTCCTCCATTTTTATTTGAGGTGTTTCTCTATGCTCCCATAGATGATGGG 183

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |  |||||||| |||||||| |||||||| ||| ||||||| |||||||||| ||  |
Db        184 AAGAGCAAAGCCAAGATTGGAAATGACACGATTCTCTTCACATTGTATAAAAAGGAGCCA 243

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |   ||||||||| ||||||| || | || ||||| |||||||||||||| ||||| |||
Db        244 GTTCTGTGGGAGAGCCTTTCTATGCCAGGCGTTGATAAAGAGATGATGCAGAGAATAAGA 303

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              |||||||||| || |||||||| |||| |||||||| |||||||||| ||||||||
Db        304 GAAAAATCTATCTTGCAAGCACAGGAGAAAGCAAAAGAGGCTACAGAAGCGAAAGCTGCT 363

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              || ||||| |||||||| | |||||| ||| | |  |||||||||||||||||| ||||||
Db        364 GCCAAGCGAGAAGATCAGAGATACGCCCTAGGCGAGATGATGAAGATTGAAGAGGAAGAG 423

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||| |||||| || | || ||| |||
Db        424 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGAAAAAAGCAACCAGAGAATTAGAA 483

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              || |||||||||  ||||  |||||||| |  |||||  | | |||||  | |||||
Db        484 GCGTGGAAAGAATGCCAAAAGAAAGCTGACGGACAAAAACGAGTCCAGAGGAAGGAGAA- 542

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
                |  | | | |||||||  |  ||||  ||   |  | |||||  | |||||
Db        543 --ACCGCTACAGGGAAAGCAAGCT---GAAGAGAGGGGAGCACTAAAACCTCAGAGTTTG 597

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              |  | ||     |  | |     ||  |  ||| |||||  ||||  |||||  |||||
Db        598 CCCCGGAAGGCCCCGCCCACTCGCCTCCCCACAAGAGGGAGGAATTGGGAAAACATATTC 657

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              |||||||||||||||||||||  | ||||| |||||||||| ||||||  ||| | |||
Db        658 TCTGAGAAGTTAAAGGAAGACAGGGTCCCTGCACCTCGCTCTGCTGGCAGCATTCAGATC 717

Qy        721 AACTTTACCCCTCGAGTATTCCCAAC 746
              | ||||  ||||||||| |||||| |
Db        718 AGCTTTNACCCTCGAGTGTTCCCACC 743


RESULT 14
CB297042
LOCUS       CB297042                468 bp    mRNA    linear   EST 28-FEB-2003
DEFINITION  12B22058_rev_1_F09_r_075.ab1 Chimpanzee brain library Koos Pan
            troglodytes cDNA clone 12B22058_rev_1_F09_r_075.ab1 5', mRNA
            sequence.
ACCESSION   CB297042
VERSION     CB297042.1  GI:28622472
KEYWORDS    EST.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 468)
  AUTHORS   Hellmann,I., Zollner,S., Enard,W., Ebersberger,I., Nickel,B. and
            Paabo,S.
  TITLE     Selection on human genes as revealed by comparisons to chimpanzee
            cDNA
  JOURNAL   Genome Res. (2003) In press
COMMENT     Contact: Paabo S
            Evolutionary Genetics
            Max-Planck-Institute for evolutionary Anthropology
            Deutscher Platz 6, 04103 Leipzig, Germany
            Tel: +49-(0)-341-3550 500
            Fax: +49-(0)-341-3550 555
            Email: paabo@eva.mpg.de
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1. .468
                     /organism="Pan troglodytes"
                     /mol_type="mRNA"
```

/db_xref="taxon:9598"
/clone="12B22058_rev_1_F09_r_075.ab1"
/sex="male"
/tissue_type="brain, presumably cortex"
/dev_stage="adult"
/lab_host="Epicurian Coli (TM) XL-10-Gold"
/clone_lib="Chimpanzee brain library Koos"
/note="Vector: pUChi; Site_1: SfiI-A; Site_2: SfiI-B; The library was prepared using the SMART cDNA library construction Kit (Clontech), doing only primer extension, but not PCR amplification of the cDNA. The only deviation from the published protocoll was that we cloned the cDNA into a plasmid Vector."

ORIGIN

Query Match 34.4%; Score 434.2; DB 6; Length 468;
Best Local Similarity 99.3%; Pred. No. 5.4e-84;
Matches 436; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         30 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCCGTCTTTCTGTCT 89

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 149

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 209

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 269

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        270 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 329

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 389

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        390 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 449

Qy        421 AGGAAAAAAAATAGAAGATA 439
              ||||||||||||||||||||
Db        450 AGGAAAAAAAATAGAAGATA 468
```

RESULT 15
BF207765
LOCUS BF207765 1183 bp mRNA linear EST 06-NOV-2000
DEFINITION 601861861F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:4081622 5', mRNA sequence.
ACCESSION BF207765
VERSION BF207765.1 GI:11101351
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1183)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM943 row: i column: 15
High quality sequence stop: 451.

FEATURES Location/Qualifiers
source 1. .1183
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4081622"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_53"
/note="Organ: bladder; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 33.8%; Score 427; DB 2; Length 1183;
Best Local Similarity 89.8%; Pred. No. 2.2e-82;
Matches 495; Conservative 0; Mismatches 45; Indels 11; Gaps 3;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db         50 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGTAGACTGCGGTCTTTCTGTCT 109

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 169

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 229

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 289

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 349

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        350 GAAAAATCTATTTTACAAGCACAAGAGAGAGCCAAAGAAGCTACAGAAGCAAAAGCTGCA 409

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAA-GATTGAAGAAGAAGA 419
              ||||||||||||||||||||||||||||||||||||||||||| ||||| |||||||||
Db        410 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAACGATTGCAGAAGAAGA 469

Qy        420 GAGGAAAAAAAATAGAAGATATGA-------AAGAAAATGAACGGATAAAAGCCACTAAAG 472
              ||||| | |   | |      ||      | ||   |||||||| ||||||||||||||
Db        470 GAGGACACACCAACATAGACAGATTCGACCACGACCCTGAACGGATCAAAGCCACTAAAG 529

Qy        473 CATTGGAAGCCTGGAAAGAATAT---CAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGA 529
              |||||||||| ||| |     |     | || | |||||||||| || | | ||
Db        530 CATTGGAAGCCGGGACAACGACTATCGACCGACACGCTGAGGAGCGAACACACATCAGAG 589

Qy        530 GAGAAGAGAAA 540
                || |||| |
Db        590 ACGACGAGACA 600
```

Search completed: July 8, 2005, 04:46:16
Job time : 4610 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 01:40:53 ; Search time 245 Seconds
(without alignments)
8435.173 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	432.6	34.3	489	4	US-09-621-976-1234	Sequence 1234, Ap
	2	178.8	14.2	313	4	US-09-513-999C-30874	Sequence 30874, A
	3	162.8	12.9	164	4	US-09-513-999C-34573	Sequence 34573, A
c	4	115.8	9.2	7218	1	US-08-232-463-14	Sequence 14, Appl
	5	90.8	7.2	612	4	US-09-902-540-1357	Sequence 1357, Ap
	6	77.4	6.1	1039	4	US-09-902-540-1280	Sequence 1280, Ap
	7	77.2	6.1	614	4	US-09-902-540-1318	Sequence 1318, Ap
	8	73.4	5.8	258775	4	US-09-949-016-16435	Sequence 16435, A
	9	72.6	5.7	705	4	US-09-270-767-5061	Sequence 5061, Ap
	10	72.6	5.7	705	4	US-09-270-767-20343	Sequence 20343, A
	11	72.4	5.7	240	1	US-08-628-417-6	Sequence 6, Appli
c	12	72	5.7	396	4	US-09-640-173-53	Sequence 53, Appl
c	13	72	5.7	396	4	US-09-713-550-53	Sequence 53, Appl
c	14	72	5.7	396	4	US-09-825-294-53	Sequence 53, Appl
c	15	72	5.7	396	4	US-09-970-966-53	Sequence 53, Appl
	16	71.2	5.6	1696	4	US-09-835-811-1	Sequence 1, Appli
	17	71.2	5.6	1827	4	US-09-270-767-1308	Sequence 1308, Ap
	18	71.2	5.6	1827	4	US-09-270-767-16590	Sequence 16590, A
c	19	70.2	5.6	851	4	US-09-270-767-8286	Sequence 8286, Ap
c	20	70.2	5.6	851	4	US-09-270-767-23568	Sequence 23568, A
	21	69.8	5.5	161124	4	US-09-949-016-11760	Sequence 11760, A
	22	69.4	5.5	2447	2	US-09-014-969-14	Sequence 14, Appl
c	23	67.4	5.3	25431	4	US-09-949-016-13234	Sequence 13234, A
	24	67.2	5.3	601	4	US-09-949-016-184917	Sequence 184917,
	25	67.2	5.3	601	4	US-09-949-016-185072	Sequence 185072,
	26	66.6	5.3	194937	4	US-09-949-016-17032	Sequence 17032, A
	27	66.6	5.3	194937	4	US-09-949-016-17033	Sequence 17033, A
	28	66.2	5.2	696	3	US-09-461-697-193	Sequence 193, App
	29	66.2	5.2	699	3	US-09-461-697-191	Sequence 191, App
	30	66.2	5.2	717	3	US-09-461-697-189	Sequence 189, App
	31	66.2	5.2	774	3	US-09-461-697-187	Sequence 187, App
	32	66.2	5.2	819	3	US-09-461-697-185	Sequence 185, App
	33	66.2	5.2	1669	3	US-09-461-697-184	Sequence 184, App
	34	66.2	5.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
	35	65.4	5.2	36731	4	US-09-949-016-13770	Sequence 13770, A
	36	64.6	5.1	2223	1	US-08-257-073-4	Sequence 4, Appli
	37	63.6	5.0	980	3	US-09-171-209-8	Sequence 8, Appli
c	38	63.6	5.0	55841	4	US-09-949-016-16602	Sequence 16602, A
	39	63.6	5.0	74790	4	US-09-949-016-15321	Sequence 15321, A
	40	63.4	5.0	5433	3	US-08-929-329-1	Sequence 1, Appli
c	41	63.4	5.0	12703	4	US-09-949-016-16685	Sequence 16685, A
	42	62.6	5.0	237510	4	US-09-949-016-14273	Sequence 14273, A
c	43	62.4	4.9	54033	4	US-09-949-016-12091	Sequence 12091, A
c	44	62.4	4.9	54033	4	US-09-949-016-14325	Sequence 14325, A
	45	62	4.9	1298	3	US-08-948-705-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-621-976-1234
; Sequence 1234, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1234
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..488
US-09-621-976-1234

Query Match 34.3%; Score 432.6; DB 4; Length 489;
Best Local Similarity 99.8%; Pred. No. 4.6e-97;
Matches 432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||
Db         57 ATGCCTCTTCAGGTTAGCGATTACAKCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 116

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 176

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 236

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 296

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 356

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 416
```

```
Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 476

Qy        421 AGGAAAAAAATAG 433
              |||||||||||||
Db        477 AGGAAAAAAATAG 489


RESULT 2
US-09-513-999C-30874
; Sequence 30874, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 30874
;   LENGTH: 313
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 82
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 84
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 85
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 89
;   OTHER INFORMATION: v=a or c or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 92
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 250
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 283
;   OTHER INFORMATION: s=g or c
US-09-513-999C-30874

  Query Match             14.2%;  Score 178.8;  DB 4;  Length 313;
  Best Local Similarity   94.8%;  Pred. No. 1.5e-34;
  Matches  201;  Conservative    5;  Mismatches    3;  Indels    3;  Gaps    2;

Qy        844 GATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTT 903
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTT 60

Qy        904 GCAACGGAAAACTATTTGGCAGCTA--TCAATGCATAT-AATTTAGCCATAAGACTAAAT 960
              |||||||||||||||||||||:|::   :||:|||||| ||||||||||||||||||||
Db         61 GCAACGGAAAACTATTTGGCARCWWATCVAAWGCATATAAATTTAGCCATAAGACTAAAT 120

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 180

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACT 1052
              |||||||||||||||||||||||||| | ||
Db        181 CACAAGGCTATTGAAGATTCTTCTAAGGAATT 212


RESULT 3
US-09-513-999C-34573
; Sequence 34573, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 34573
;   LENGTH: 164
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 127
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 128
;   OTHER INFORMATION: h=a or c or t
US-09-513-999C-34573

  Query Match             12.9%;  Score 162.8;  DB 4;  Length 164;
  Best Local Similarity   98.8%;  Pred. No. 9.9e-31;
  Matches  162;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 CTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGC 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGC 60

Qy         65 CCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGG 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGG 120

Qy        125 TCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 168
              ||||||::|||||||||||||||||||||||||||||||||||||
Db        121 TCAACTYHCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 164


RESULT 4
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
```

```
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             9.2%;  Score 115.8;  DB 1;  Length 7218;
  Best Local Similarity   7.0%;  Pred. No. 1.7e-18;
  Matches   30;  Conservative  271;  Mismatches  128;  Indels    0;  Gaps    0;

Qy        298 AGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCT 357
              | |||| ||| | |  : ::: : :::::::::: ::::::::  : ::::: :::::
Db       1451 ATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1392

Qy        358 GCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAA 417
              : :: :::: :::::::  ::::: : : :  ::: :  : :: :::::  :::::::::
Db       1391 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1332

Qy        418 GAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTG 477
              :::::::::::::: :::::: : :::::::::: ::: ::: ::::::  :  :::: :  :
Db       1331 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1272

Qy        478 GAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAG 537
              ::::   :::::::: :  :::::::::  :::::: ::::::::::   ::::::::::
Db       1271 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1212

Qy        538 AAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGT 597
              :::  : :  :::::::::: :::  :::::::::::::::::: :::: : :::::
Db       1211 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1152

Qy        598 CTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATA 657
              :  :::::   :: :  :::::    :    :::::::::::   :::::: : :
Db       1151 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092

Qy        658 TTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAA 717
              :  ::::::  ::::::::: : |          |||||  |||  | ||   |   |
Db       1091 RRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGCCAAGCTCGGA 1032

Qy        718 ATCAACTTT 726
              || || | |
Db       1031 ATTAATTCT 1023


RESULT 5
US-09-902-540-1357
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
;   LENGTH: 612
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(612)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

  Query Match             7.2%;  Score 90.8;  DB 4;  Length 612;
  Best Local Similarity   51.2%;  Pred. No. 9.5e-13;
  Matches  209;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy        275 ACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAA 334
              | ||| | |  |   ||| ||  | | |||||   |    | ||  | || | | |  ||
Db         70 AAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 129

Qy        335 AAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTG 394
              || ||   | | ||  ||||     |  |||     || |  ||||  |   |  ||
Db        130 AAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAA 189

Qy        395 TCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAAC 454
                |  |  || ||  || || || | |  ||||||||| || | |  ||| ||||   |
Db        190 AAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAATAA 249

Qy        455 GGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGC 514
                | ||||   |  | |  | |  ||     ||| | ||  ||| ||     |   |
Db        250 AAAAAAAAAATAAAATAAAAATAAAAAAAAATAAAATATTAAAAAAAAAATTAAAAAAAAAA 309

Qy        515 AAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAA 574
              |||  |||||  | | | || | || | |    | ||| ||||   || |||| ||  ||
Db        310 AAANTAAAATAAAAAAAAAAAAAAAANTAAAAAAATAAAAAAAAAATAAATAAAAAAAAAA 369

Qy        575 GAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAA 634
              |||||||| |||||| || |    |  | |||       |  || |||           ||
Db        370 TAAAAAAAAAAAAAATAAAATATAAAAATAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAA 429

Qy        635 AAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACA 682
              ||   | ||||| | |||| ||    |   |  |||  || | |
Db        430 AAAAAAAAAATTAAAAAAAAATTAAAAAATAATAAAAAAAAAAAAATA 477


RESULT 6
US-09-902-540-1280
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
```

```
;   LENGTH: 1039
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(1039)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

  Query Match             6.1%;  Score 77.4;  DB 4;  Length 1039;
  Best Local Similarity   50.5%;  Pred. No. 2.4e-09;
  Matches  186;  Conservative    0;  Mismatches  182;  Indels    0;  Gaps    0;

Qy        290 AAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAG 349
              ||  ||  | | |||||    |    | ||  | || | |     ||| ||   | | ||
Db         79 AACAAAAAATAAAAAAAAAAAAAAAATAAAAAAAACAAATTAATAAAAAAAAAATAAAAA 138

Qy        350 CAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTG 409
              ||||       |  |||      || |  |  || |    |  |       |  |  || |
Db        139 CAAATATAAAAAAAAAAAAAAAAAAAATATTAAAAAAAAAAAATAAAAAAAAAAAAAAAAA 198

Qy        410 AAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTA 469
              || || || | |  || |  ||| || | |  ||| |||| |    | ||||  || ||
Db        199 AAAAAAAAAAAAAAAACATTATAAAAAAAAAAAAAAAAAAAAATATATAAAAAAAAACAATA 258

Qy        470 AAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGA 529
              |  |    ||      ||| |  |  ||| |||| |  |  |  |||||| ||   | |
Db        259 CATAAAAAAAAAAAAAAAAAAATAAATAAAAAAAATCAAAAAAAAAAAAAAATAAAAAAAA 318

Qy        530 GAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAAT 589
              | ||   ||| |   ||| |||  |||  |||||||| ||  || |||||||| ||||
Db        319 AAAAAATTAAAATCAATCATAAAATAAAAAAAATTAAATAAATAACAAAAAAAACAAAAA 378

Qy        590 ATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAG 649
              || | |  |   | |  |||||   |  |  ||        ||||   | |||   |
Db        379 ATTAAAATAAAAAAATTAATTTACAACAAAATAAAAGAAAANTAAAAAAAAAAAAAAAAT 438

Qy        650 AAAATATA 657
              || || ||
Db        439 AATATGTA 446


RESULT 7
US-09-902-540-1318
; Sequence 1318, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1318
;   LENGTH: 614
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (1)..(614)
;   OTHER INFORMATION: unsure at all n locations
US-09-902-540-1318

  Query Match             6.1%;  Score 77.2;  DB 4;  Length 614;
  Best Local Similarity   52.9%;  Pred. No. 2.1e-09;
  Matches  166;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy        275 ACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAA 334
              | ||  | |  | | ||| ||  ||| | | |    |    ||||    || | | |  ||
Db        290 AAAACAACAAAAAGAAAAAAAAAAGAAAGACAAAAAAAAAAACAAAACAAAAACAAAATAA 349

Qy        335 AAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTG 394
              || ||    |   ||  | |  |  ||  ||| |   ||    ||||| | | |  ||
Db        350 AAAAAAAAAACAAAAAACACACAACAAAAAAACATAAAACAAAAAAAAAAGAAAAAAAAA 409

Qy        395 TCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAAC 454
              |  | || |   || || || | |  ||||||| |  | | |  ||| ||||  ||
Db        410 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAAA 469

Qy        455 GGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGC 514
                | |||||  |  |||  |    ||      ||| ||||  ||| ||||    |  |
Db        470 ACAAAAAAGAAAAAAAAAAAAACAAACAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAA 529

Qy        515 AAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAA 574
              |||||||||      ||| | ||| |  | |  ||| | | ||  |||  | | ||  ||
Db        530 AAAAAAAAAACACAAGAAGAAAGACAAAAAGAAAAAGAAACAAAAAAAAAAGAAAATCAA 589

Qy        575 GAAAAAAAATAAAA 588
              ||||| || ||||
Db        590 AAAAAAGAAAAAAA 603


RESULT 8
US-09-949-016-16435
; Sequence 16435, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16435
;   LENGTH: 258775
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16435

  Query Match             5.8%;  Score 73.4;  DB 4;  Length 258775;
  Best Local Similarity   51.7%;  Pred. No. 2.1e-07;
  Matches  167;  Conservative    0;  Mismatches  156;  Indels    0;  Gaps    0;

Qy        274 GACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCA 333
              |||| | ||| ||   ||||||  | ||||| |    |    | |    |||| |||
Db     122487 GACACATAGACGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 122546

Qy        334 AAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGT 393
              ||||||    |   ||| || |          | ||    ||||   |||  | | |  ||
Db     122547 AAAGAAAGAAAGAAAGAAAGAAAGAGGAGAGAGAAAAAAGAAAGAAAGAAAGAAAGAAAG 122606

Qy        394 GTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAA 453
                  |  ||||  |||  | | |||||  ||||||  ||||    | |||||||    |
Db     122607 AAAGAAAGAAAGAAAGAAAGAAAGGAGAGAGAAAAAAGAGAAAGAAAGAAAGAAAGAAAG 122666

Qy        454 CGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAG 513
                    ||||  |  ||   |      ||    || | || |   ||  ||||      |||
Db     122667 AAAGAGAAAGAAAGAAAGAAAGAAAGAGGAGAGAGAAAAAAGAGAAAGAAAGAAAGAGAG 122726

Qy        514 CAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGA 573
              |||  | |||    |||   |||| | |          |||  || |   ||  |||||| ||
Db     122727 AAAAGAGAAAGAAAGAAGAAAGAAAGAAAGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGA 122786
```

```
Qy        574 AGAAAAAAAATAAAATATAAGAG 596
              |  ||| ||| |||  |  | ||
Db     122787 AAGAAAGAAAGAAAGAAAGAAAG 122809
```

RESULT 9
US-09-270-767-5061
; Sequence 5061, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5061
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-5061

Query Match 5.7%; Score 72.6; DB 4; Length 705;
Best Local Similarity 51.8%; Pred. No. 3.1e-08;
Matches 216; Conservative 0; Mismatches 194; Indels 7; Gaps 2;

```
Qy        264 GACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACA 323
              || | |   |  ||||| | ||   ||   |  ||| |||        | |||   |
Db        171 GAAGAGAAGAGGGAAAGAAAAGAGAGAAGAGAAAAGAAGAAAGAGAGGGGGAAAAGAGGA 230

Qy        324 AGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATA 383
              |||||     |||||||      |    ||||  | ||  |   || || |  |||   |
Db        231 AGAGAAGAGGAAAGAAGAAGGGAAGAGGAAAGAAGAAGAGAGAAGGAAAAAAGAAAGGAA 290

Qy        384 CGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAA 443
                  | || || ||  ||| ||||  ||||||||||  | |  | | ||
Db        291 GAGGAACGAAGAAAAGAAGAGGAGAGAAAAAGAGAAGAGGAAAAAAGAAAAGAAGAGGAG 350

Qy        444 AGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAA 503
              ||||||  ||   ||  ||||  |  | || |  ||||     | || ||| |  || |||
Db        351 AGAAAAAGAGAAGAGGAAAG-AAGAAGAGAAGAGGAAAAAAGAAAGGAAGAGGAACGAAG 409

Qy        504 AGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAAT 563
              |    || ||| | | ||||      |||||||||  ||  | |  || |||| |||  ||
Db        410 AAAAGAAGAGGAGAGAAAA------AGAGAAGAGGAAAAAAGAAAAGAAGAGAAGAAA 463

Qy        564 TAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAA 623
              | |||||     |||||  || |||| |  ||||     |     | |        |   | | |
Db        464 AAGAGAAGAGGAAAAAAGAAAAAAGAGGAGAGAAAAAGAGAAGAGGAAAAAAGAAAAGA 523

Qy        624 TCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGA 680
                ||    |||| |   | |   | |||  | |      || ||   |||||| ||
Db        524 AGATGAGAGAAAAAGAGAAGAGGAAAAAAGAAAAGAAGAGGAAAAAAGAAAGGAGGA 580

Query Match 5.7%; Score 72.6; DB 4; Length 705;
Best Local Similarity 51.8%; Pred. No. 3.1e-08;
Matches 216; Conservative 0; Mismatches 194; Indels 7; Gaps 2;

```
Qy        264 GACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACA 323
              || | |   |  ||||| | ||   ||   |  ||| |||        | |||   |
Db        171 GAAGAGAAGAGGGAAAGAAAAGAGAGAAGAGAAAAGAAGAAAGAGAGGGGGAAAAGAGGA 230

Qy        324 AGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATA 383
              |||||     |||||||      |    ||||  | ||  |   || || |  |||   |
Db        231 AGAGAAGAGGAAAGAAGAAGGGAAGAGGAAAGAAGAAGAGAGAAGGAAAAAAGAAAGGAA 290

Qy        384 CGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAA 443
                  | || || ||  ||| ||||  ||||||||||  | |  | | ||
Db        291 GAGGAACGAAGAAAAGAAGAGGAGAGAAAAAGAGAAGAGGAAAAAAGAAAAGAAGAGGAG 350

Qy        444 AGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAA 503
              ||||||  ||   ||  ||||  |  | || |  ||||     | || ||| |  || |||
Db        351 AGAAAAAGAGAAGAGGAAAG-AAGAAGAGAAGAGGAAAAAAGAAAGGAAGAGGAACGAAG 409

Qy        504 AGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAAT 563
              |    || ||| | | ||||      |||||||||  ||  | |  || |||| |||  ||
Db        410 AAAAGAAGAGGAGAGAAAA------AGAGAAGAGGAAAAAAGAAAAGAAGAGAAGAAA 463

Qy        564 TAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAA 623
              | |||||     |||||  || |||| |  ||||     |     | |        |   | | |
Db        464 AAGAGAAGAGGAAAAAAGAAAAAAGAGGAGAGAAAAAGAGAAGAGGAAAAAAGAAAAGA 523

Qy        624 TCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGA 680
                ||    |||| |   | |   | |||  | |      || ||   |||||| ||
Db        524 AGATGAGAGAAAAAGAGAAGAGGAAAAAAGAAAAGAAGAGGAAAAAAGAAAGGAGGA 580
```

RESULT 10
US-09-270-767-20343
; Sequence 20343, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20343
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-20343

RESULT 11
US-08-628-417-6
; Sequence 6, Application US/08628417
; Patent No. 5627054
; GENERAL INFORMATION:
; APPLICANT: GILLESPIE, DAVID
; TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC
; TITLE OF INVENTION: POLYMERASE CHAIN REACTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL
; ADDRESSEE: DEFENSE COMMAND
; STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)
; CITY: ABERDEEN PROVING GROUND
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,417
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BIFFONI, ULYSSES J
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 398-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
;    MOLECULE TYPE:  oligodeoxynucleotide
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-628-417-6

  Query Match             5.7%;  Score 72.4;  DB 1;  Length 240;
  Best Local Similarity   57.5%;  Pred. No. 2.2e-08;
  Matches  130;  Conservative    0;  Mismatches   96;  Indels    0;  Gaps    0;

Qy        370 GAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAA 429
              ||  || |||  |  | | ||| | |  | |  || |   || || || | |  ||||||
Db         11 GACAATAAAACTTTAGAAATAATTTTACTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 70

Qy        430 ATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAA 489
              | | || | |  ||| ||||  ||   | ||||   |  |||  |    ||       |||
Db         71 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 130

Qy        490 GAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAA 549
               || |  ||| ||||    |  |  |||||||||   | | | || | |||  |    ||
Db        131 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 190

Qy        550 AAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGA 595
              ||| ||||  |||  ||| ||  || |||||||| |||| | || |
Db        191 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 236


RESULT 12
US-09-640-173-53/c
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION:  OVARIAN TUMOR SEQUENCES AND
;  TITLE OF INVENTION:  METHODS OF USE THEREFOR
;  FILE REFERENCE:  210121.484C2
;  CURRENT APPLICATION NUMBER: US/09/640,173
;  CURRENT FILING DATE:  2000-08-15
;  NUMBER OF SEQ ID NOS:  196
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 53
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

  Query Match             5.7%;  Score 72;  DB 4;  Length 396;
  Best Local Similarity   54.9%;  Pred. No. 3.4e-08;
  Matches  135;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy        363 AAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAG 422
              | |  || || |  ||||| |       ||      |  |  || |    || || || | |
Db        247 ANAAAGGNAAAANAAAAAANAANNTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188

Qy        423 GAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGC 482
              |||||||  | || | |  ||| ||||  ||   | ||||   |  |||  |    ||
Db        187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

Qy        483 CTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATT 542
                 ||| || |  ||| ||||    |  |  |||||||||   | | | || | |||
Db        127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68

Qy        543 ATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTAC 602
              |    ||||| ||||  |||  ||| ||  || |||||||| |||| | || |     |
Db         67 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

Qy        603 TAGAAA 608
                | |||
Db          7 AAAAAA 2


RESULT 13
US-09-713-550-53/c
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.484C4
;  CURRENT APPLICATION NUMBER: US/09/713,550
;  CURRENT FILING DATE:  2000-11-14
;  NUMBER OF SEQ ID NOS:  205
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 53
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

  Query Match             5.7%;  Score 72;  DB 4;  Length 396;
  Best Local Similarity   54.9%;  Pred. No. 3.4e-08;
  Matches  135;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy        363 AAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAG 422
              | |  || || |  ||||| |       ||      |  |  || |    || || || | |
Db        247 ANAAAGGNAAAANAAAAAANAANNTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188

Qy        423 GAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGC 482
              |||||||  | || | |  ||| ||||  ||   | ||||   |  |||  |    ||
Db        187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

Qy        483 CTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATT 542
                 ||| || |  ||| ||||    |  |  |||||||||   | | | || | |||
Db        127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68

Qy        543 ATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTAC 602
              |    ||||| ||||  |||  ||| ||  || |||||||| |||| | || |     |
Db         67 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

Qy        603 TAGAAA 608
                | |||
Db          7 AAAAAA 2


RESULT 14
US-09-825-294-53/c
; Sequence 53, Application US/09825294
; Patent No. 6710170
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  APPLICANT:  Stolk, John A.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.484C5
;  CURRENT APPLICATION NUMBER: US/09/825,294
;  CURRENT FILING DATE:  2001-04-03
;  NUMBER OF SEQ ID NOS: 215
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 53
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (1)...(396)
;   OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53

  Query Match             5.7%;  Score 72;  DB 4;  Length 396;
  Best Local Similarity   54.9%;  Pred. No. 3.4e-08;
  Matches  135;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy        363 AAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAG 422
              | |  || || |  ||||| |       ||      |  |  || |   || || || | |
Db        247 ANAAAGGNAAAANAAAAAANAANNTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188

Qy        423 GAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGC 482
               ||||||| | || | |  ||| ||||  ||   | ||||   |  |||  |    ||
Db        187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

Qy        483 CTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATT 542
              |||  || |  ||| ||||    |  |  |||||||||   | | |  || | |||
Db        127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68

Qy        543 ATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTAC 602
              |    ||||| ||||  |||  ||| ||  || |||||||| |||| | || |      |
Db         67 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

Qy        603 TAGAAA 608
                | |||
Db          7 AAAAAA 2


RESULT 15
US-09-970-966-53/c
; Sequence 53, Application US/09970966
; Patent No. 6720146
; GENERAL INFORMATION:
;  APPLICANT: Stolk, John A.
;  APPLICANT:  Molesh, David Alan
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Xu, Jiangchun
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.484C6
;  CURRENT APPLICATION NUMBER: US/09/970,966
;  CURRENT FILING DATE:  2001-10-02
;  NUMBER OF SEQ ID NOS: 215
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 224, 225, 228, 235, 240, 246, 257, 266, 274, 279, 281, 282,
;   LOCATION: 283, 285, 287, 288, 290, 291, 292, 293, 294, 295, 296, 297,
;   LOCATION: 300, 301, 303, 307, 311, 313, 314, 317, 318, 319, 320, 321,
;   LOCATION: 323, 324, 328, 329, 330, 336, 337, 338, 339, 340, 341
;   OTHER INFORMATION: n = A,T,C or G
;   NAME/KEY: misc_feature
;   LOCATION: 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 356,
;   LOCATION: 357, 358, 359, 362, 363, 364, 365, 366, 367, 373, 380, 381,
;   LOCATION: 382, 385, 387, 388, 389, 390, 392
;   OTHER INFORMATION: n = A,T,C or G
US-09-970-966-53

  Query Match             5.7%;  Score 72;  DB 4;  Length 396;
  Best Local Similarity   54.9%;  Pred. No. 3.4e-08;
  Matches  135;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy        363 AAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAG 422
              | |  || || |  ||||| |       ||      |  |  || |   || || || | |
Db        247 ANAAAGGNAAAANAAAAAANAANNTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188

Qy        423 GAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGC 482
               ||||||| | || | |  ||| ||||  ||   | ||||   |  |||  |    ||
Db        187 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

Qy        483 CTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATT 542
              |||  || |  ||| ||||    |  |  |||||||||   | | |  || | |||
Db        127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68

Qy        543 ATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTAC 602
              |    ||||| ||||  |||  ||| ||  || |||||||| |||| | || |      |
Db         67 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8

Qy        603 TAGAAA 608
                | |||
Db          7 AAAAAA 2


Search completed: July  8, 2005, 04:50:19
Job time : 248 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 03:29:26 ; Search time 1941 Seconds
(without alignments)
4085.271 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1263	100.0	1263	17	US-10-364-505-1	Sequence 1, Appli
2	1263	100.0	1263	19	US-10-681-199-1	Sequence 1, Appli
3	1263	100.0	1993	17	US-10-364-505-2	Sequence 2, Appli
4	1263	100.0	1993	19	US-10-681-199-2	Sequence 2, Appli
5	1253.4	99.2	1263	19	US-10-681-199-19	Sequence 19, Appl
6	1251.8	99.1	1263	19	US-10-681-199-13	Sequence 13, Appl
7	1248.6	98.9	1263	19	US-10-681-199-15	Sequence 15, Appl
8	1243.8	98.5	1263	19	US-10-681-199-17	Sequence 17, Appl
9	1047.4	82.9	1559	17	US-10-108-260A-575	Sequence 575, App
10	840.8	66.6	1697	17	US-10-364-505-4	Sequence 4, Appli
11	840.8	66.6	1697	19	US-10-681-199-4	Sequence 4, Appli
12	401.6	31.8	464	17	US-10-242-535A-30001	Sequence 30001, A
13	401.6	31.8	464	18	US-10-085-783A-30001	Sequence 30001, A
14	369.8	29.3	488	10	US-09-918-995-22675	Sequence 22675, A
15	336	26.6	458	10	US-09-918-995-10303	Sequence 10303, A
c 16	325	25.7	325	18	US-10-240-425-872	Sequence 872, App
c 17	264.2	20.9	715	21	US-10-956-157-3114	Sequence 3114, Ap
18	264.2	20.9	715	21	US-10-956-157-8349	Sequence 8349, Ap
19	239.4	19.0	50000	17	US-10-364-505-7	Sequence 7, Appli
20	239.4	19.0	50000	19	US-10-681-199-7	Sequence 7, Appli
21	219.2	17.4	505	9	US-09-783-590-8049	Sequence 8049, Ap
22	156.6	12.4	50000	17	US-10-364-505-8	Sequence 8, Appli
23	156.6	12.4	50000	19	US-10-681-199-8	Sequence 8, Appli
c 24	99	7.8	1204	19	US-10-437-963-77858	Sequence 77858, A
c 25	97.6	7.7	14006	15	US-10-311-455-1931	Sequence 1931, Ap
c 26	97.6	7.7	3673778	16	US-10-312-841-2	Sequence 2, Appli
c 27	96.8	7.7	16033	15	US-10-311-455-1377	Sequence 1377, Ap
c 28	95.4	7.6	803	19	US-10-437-963-72176	Sequence 72176, A
c 29	94.8	7.5	880	20	US-10-425-115-53133	Sequence 53133, A
c 30	94.2	7.5	1243	20	US-10-425-115-172717	Sequence 172717,
c 31	94	7.4	921	20	US-10-425-115-38710	Sequence 38710, A
32	93.4	7.4	447	19	US-10-021-323-9573	Sequence 9573, Ap
c 33	93.4	7.4	986	20	US-10-425-115-174478	Sequence 174478,
c 34	93.2	7.4	529	9	US-09-983-965-2109	Sequence 2109, Ap
35	93.2	7.4	2520	17	US-10-389-566-205	Sequence 205, App
c 36	93	7.4	1062	20	US-10-425-115-120013	Sequence 120013,
37	92.8	7.3	627	19	US-10-021-323-9336	Sequence 9336, Ap
c 38	92.8	7.3	6668	15	US-10-311-455-1670	Sequence 1670, Ap
c 39	92.8	7.3	17934	15	US-10-311-455-1692	Sequence 1692, Ap
c 40	92.4	7.3	602	20	US-10-425-115-182618	Sequence 182618,
c 41	92.2	7.3	556	20	US-10-363-345A-23587	Sequence 23587, A
42	92.2	7.3	556	20	US-10-363-345A-23588	Sequence 23588, A
c 43	92.2	7.3	556	21	US-10-363-483A-23587	Sequence 23587, A
44	92.2	7.3	556	21	US-10-363-483A-23588	Sequence 23588, A
45	91.4	7.2	1214	18	US-10-424-599-102083	Sequence 102083,

ALIGNMENTS

RESULT 1
US-10-364-505-1
; Sequence 1, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
; APPLICANT: Kere, Juha
; APPLICANT: Taipale, Mikko
; APPLICANT: No. US20030219787A1ola-Hemmi, Jaana
; APPLICANT: Kaminen, Nina
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0199P
; CURRENT APPLICATION NUMBER: US/10/364,505
; CURRENT FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequence for human DYXC1 (cDNA)
US-10-364-505-1

Query Match 100.0%; Score 1263; DB 17; Length 1263;
Best Local Similarity 100.0%; Pred. No. 5e-253;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

```
Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263
```

RESULT 2
US-10-681-199-1
; Sequence 1, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequence for human DYXC1 (cDNA)
US-10-681-199-1

Query Match 100.0%; Score 1263; DB 19; Length 1263;
Best Local Similarity 100.0%; Pred. No. 5e-253;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
```

```
Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263
```

```
RESULT 3
US-10-364-505-2
; Sequence 2, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
;  APPLICANT: Kere, Juha
;  APPLICANT:  Taipale, Mikko
;  APPLICANT:  No. US20030219787A1ola-Hemmi, Jaana
;  APPLICANT:  Kaminen, Nina
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0199P
;  CURRENT APPLICATION NUMBER: US/10/364,505
;  CURRENT FILING DATE:  2003-02-12
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1993
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (369)..(1628)
;   FEATURE:
;   OTHER INFORMATION: human DYXC1 mRNA as cDNA
US-10-364-505-2

  Query Match            100.0%;  Score 1263;  DB 17;  Length 1993;
  Best Local Similarity  100.0%;  Pred. No. 5.9e-253;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328
```

```
Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631


RESULT 4
US-10-681-199-2
; Sequence 2, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1993
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (369)..(1628)
;   FEATURE:
;   OTHER INFORMATION: human DYXC1 mRNA as cDNA
US-10-681-199-2

  Query Match             100.0%;  Score 1263;  DB 19;  Length 1993;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-253;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631


RESULT 5
US-10-681-199-19
; Sequence 19, Application US/10681199
; Publication No. US20040138441A1
```

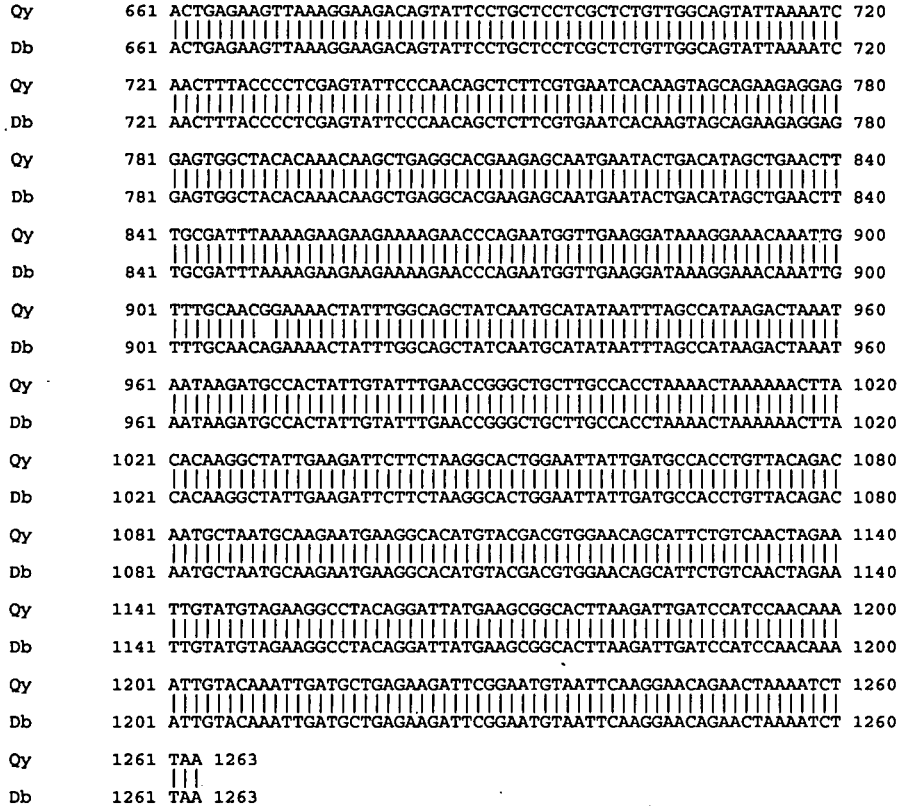
```
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 1263
;   TYPE: DNA
;   ORGANISM: Pan paniscus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1260)
US-10-681-199-19

  Query Match             99.2%;  Score 1253.4;  DB 19;  Length 1263;
  Best Local Similarity   99.5%;  Pred. No. 5e-251;
  Matches 1257;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTGCACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263


RESULT 6
US-10-681-199-13
; Sequence 13, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 1263
;   TYPE: DNA
;   ORGANISM: Pan troglodytes
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1260)
US-10-681-199-13

  Query Match             99.1%;  Score 1251.8;  DB 19;  Length 1263;
  Best Local Similarity   99.4%;  Pred. No. 1.1e-250;
  Matches 1256;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCCGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCATGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
```

```
Qy       181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
Db       181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy       241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
Db       241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy       301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
Db       301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy       361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
Db       361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy       421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
Db       421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy       481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
Db       481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy       541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
Db       541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTT 600

Qy       601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
Db       601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy       661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
Db       661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy       721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
Db       721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy       781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
Db       781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy       841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
Db       841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy       901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
Db       901 TTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy       961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
Db       961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy      1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
Db      1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy      1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
Db      1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy      1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
Db      1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy      1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
Db      1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy      1261 TAA 1263
Db      1261 TAA 1263
```

RESULT 7
US-10-681-199-15
; Sequence 15, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Gorilla gorilla
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-15

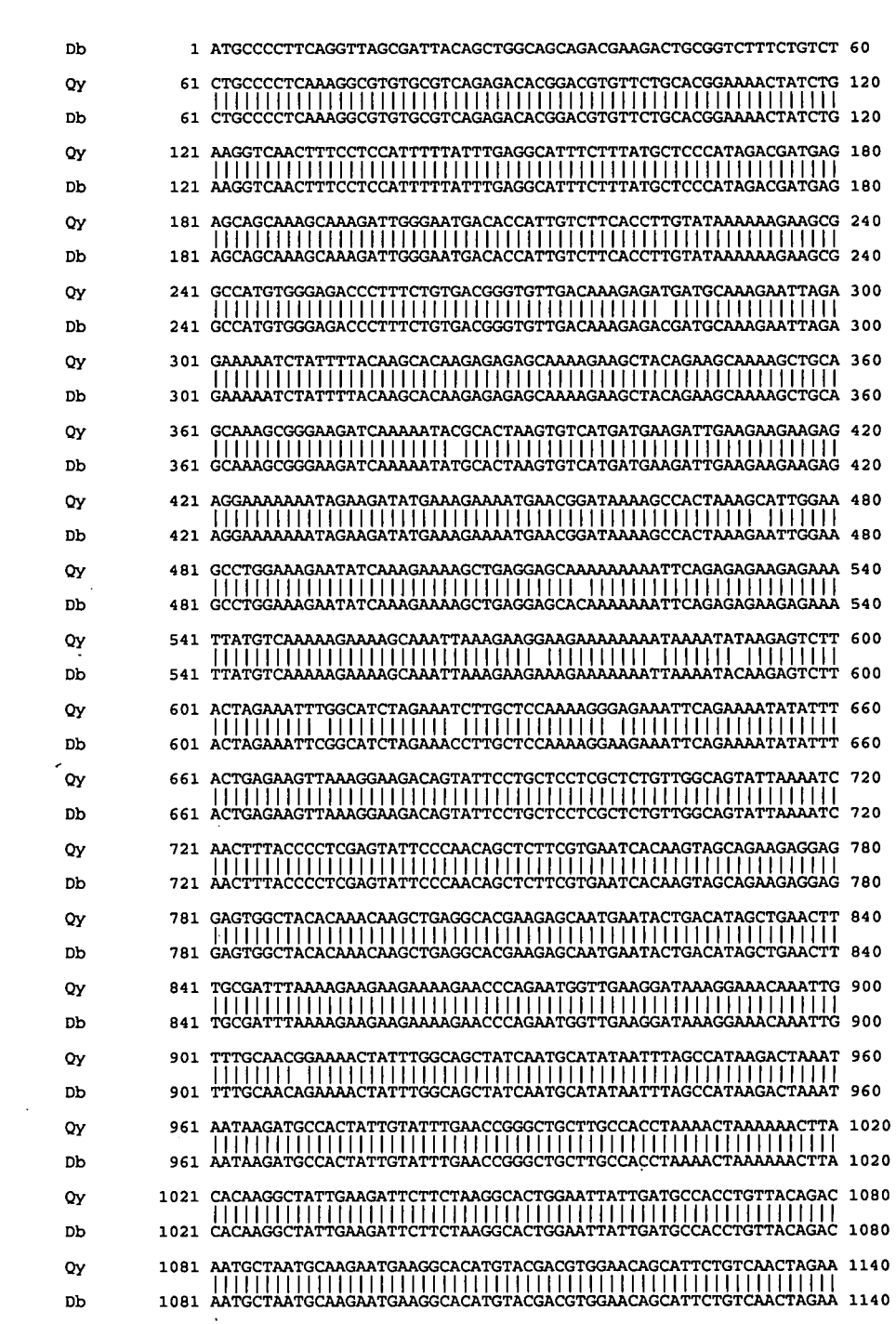
Query Match 98.9%; Score 1248.6; DB 19; Length 1263;
Best Local Similarity 99.3%; Pred. No. 5e-250;
Matches 1254; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Qy         1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
Db         1 ATGCCCCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGTGGTCTTTCTGTCT 60

Qy        61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
Db        61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy       121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
Db       121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy       181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
Db       181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy       241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
Db       241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy       301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
Db       301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy       361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
Db       361 GCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy       421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
Db       421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAA 480

Qy       481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
Db       481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAGAAATTCAGAGAGAAGAGAAG 540

Qy       541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
Db       541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATTAAAATATAAGAGTCTT 600

Qy       601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
Db       601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
```

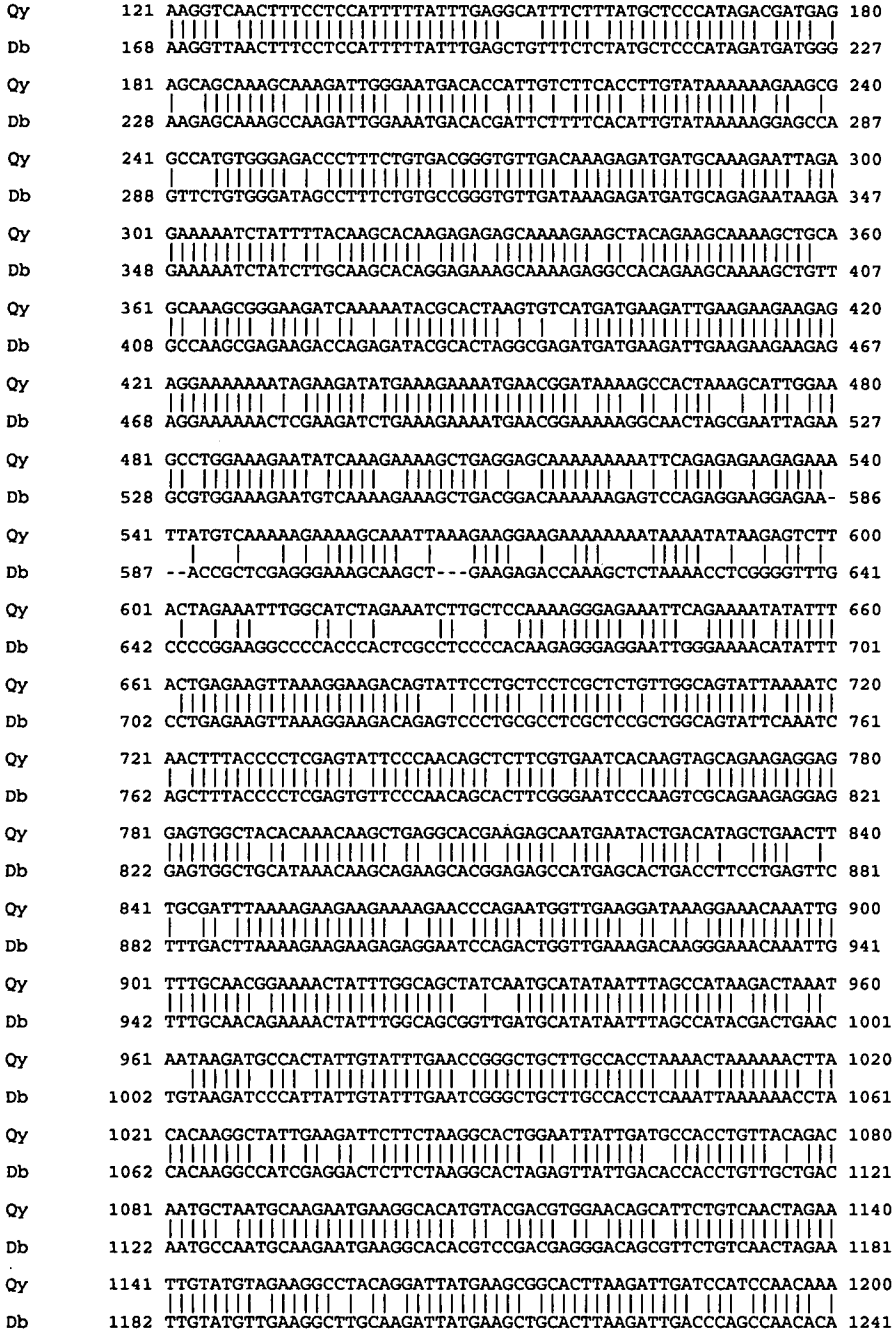


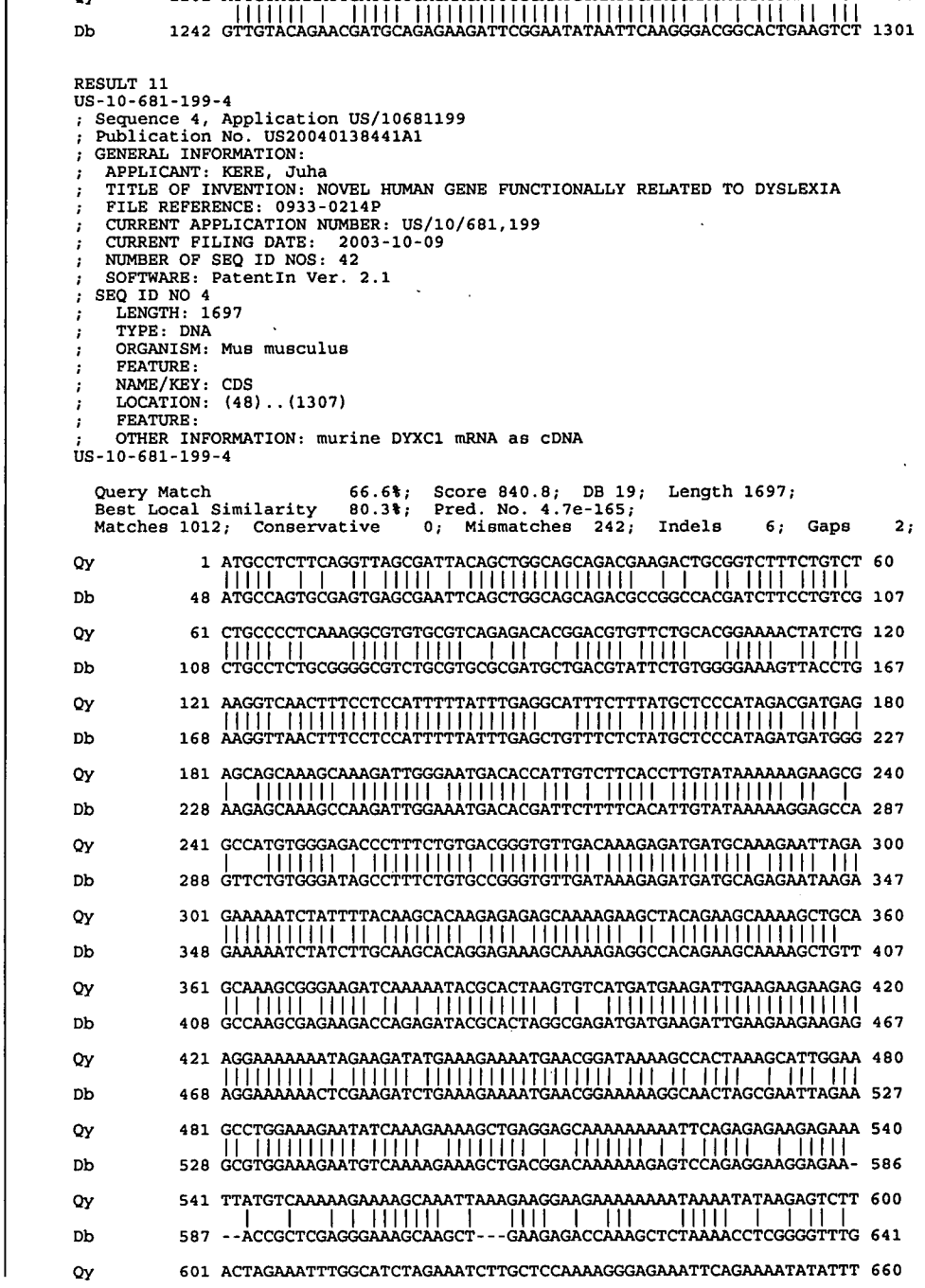
RESULT 8
US-10-681-199-17
; Sequence 17, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pongo pygmaeus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-17

Query Match 98.5%; Score 1243.8; DB 19; Length 1263;
Best Local Similarity 99.0%; Pred. No. 5e-249;
Matches 1251; Conservative 0; Mismatches 12; Indels 0; Gaps 0;



```
Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
             |||
Db       1261 TAA 1263


RESULT 9
US-10-108-260A-575
; Sequence 575, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 575
;   LENGTH: 1559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-108-260A-575

  Query Match             82.9%;  Score 1047.4;  DB 17;  Length 1559;
  Best Local Similarity   99.9%;  Pred. No. 4.3e-208;
  Matches 1048;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091


RESULT 10
US-10-364-505-4
; Sequence 4, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
;  APPLICANT: Kere, Juha
;  APPLICANT:  Taipale, Mikko
;  APPLICANT:  No. US20030219787A1ola-Hemmi, Jaana
;  APPLICANT:  Kaminen, Nina
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0199P
;  CURRENT APPLICATION NUMBER: US/10/364,505
;  CURRENT FILING DATE:  2003-02-12
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 1697
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (48)..(1307)
;   FEATURE:
;   OTHER INFORMATION: murine DYXC1 mRNA as cDNA
US-10-364-505-4

  Query Match             66.6%;  Score 840.8;  DB 17;  Length 1697;
  Best Local Similarity   80.3%;  Pred. No. 4.7e-165;
  Matches 1012;  Conservative    0;  Mismatches  242;  Indels    6;  Gaps    2;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              |||||  | |  || ||||| | ||||||||||||||||  | |  || |||| |||||
Db         48 ATGCCAGTGCGAGTGAGCGAATTCAGCTGGCAGCAGACGCCGGCCACGATCTTCCTGTCG 107

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||| ||    ||||| |||||  | ||  | ||||| |||||   |||||  || |||
Db        108 CTGCCTCTGCGGGGCGTCTGCGTGCGCGATGCTGACGTATTCTGTGGGGAAAGTTACCTG 167
```





RESULT 11
US-10-681-199-4
; Sequence 4, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1697
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(1307)
; FEATURE:
; OTHER INFORMATION: murine DYXC1 mRNA as cDNA
US-10-681-199-4

Query Match 66.6%; Score 840.8; DB 19; Length 1697;
Best Local Similarity 80.3%; Pred. No. 4.7e-165;
Matches 1012; Conservative 0; Mismatches 242; Indels 6; Gaps 2;

```
Db        642 CCCCGGAAGGCCCCACCCACTCGCCTCCCCACAAGAGGGAGGAATTGGGAAAACATATTT 701

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Db        702 CCTGAGAAGTTAAAGGAAGACAGAGTCCCTGCGCCTCGCTCCGCTGGCAGTATTCAAATC 761

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Db        762 AGCTTTACCCCTCGAGTGTTCCCAACAGCACTTCGGGAATCCCAAGTCGCAGAAGAGGAG 821

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Db        822 GAGTGGCTGCATAAACAAGCAGAAGCACGGAGAGCCATGAGCACTGACCTTCCTGAGTTC 881

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Db        882 TTTGACTTAAAAGAAGAAGAGAGGAATCCAGACTGGTTGAAAGACAAGGGAAACAAATTG 941

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Db        942 TTTGCAACAGAAAACTATTTGGCAGCGGTTGATGCATATAATTTAGCCATACGACTGAAC 1001

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Db       1002 TGTAAGATCCCATTATTGTATTTGAATCGGGCTGCTTGCCACCTCAAATTAAAAAACCTA 1061

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Db       1062 CACAAGGCCATCGAGGACTCTTCTAAGGCACTAGAGTTATTGACACCACCTGTTGCTGAC 1121

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Db       1122 AATGCCAATGCAAGAATGAAGGCACACGTCCGACGAGGGACAGCGTTCTGTCAACTAGAA 1181

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Db       1182 TTGTATGTTGAAGGCTTGCAAGATTATGAAGCTGCACTTAAGATTGACCCAGCCAACACA 1241

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Db       1242 GTTGTACAGAACGATGCAGAGAAGATTCGGAATATAATTCAAGGGACGGCACTGAAGTCT 1301
```

```
RESULT 12
US-10-242-535A-30001
; Sequence 30001, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30001
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (263)..(263)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (321)..(321)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-30001

  Query Match             31.8%;  Score 401.6;  DB 17;  Length 464;
  Best Local Similarity   98.3%;  Pred. No. 8.7e-74;
  Matches  404;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Db         49 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 108

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Db        109 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 168

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Db        169 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 228

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Db        229 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGNCTTCACCTTGTATAAAAAAGGAGCG 288

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Db        289 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTNGACAAAGAGATGATGCAAAGAATTANA 348

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Db        349 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 408

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAA 411

Db        409 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGGGAGAA 459
```

```
RESULT 13
US-10-085-783A-30001
; Sequence 30001, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30001
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (263)..(263)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (321)..(321)
```

```
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-30001

  Query Match             31.8%;  Score 401.6;  DB 18;  Length 464;
  Best Local Similarity   98.3%;  Pred. No. 8.7e-74;
  Matches  404;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 108

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 168

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 228

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||
Db        229 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGNCTTCACCTTGTATAAAAAAGGAGCG 288

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |
Db        289 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTNGACAAAGAGATGATGCAAAGAATTANA 348

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        349 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 408

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAA 411
              |||||||||||||||||||||||||||||||||||||||||||    |||
Db        409 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGGGAGAA 459


RESULT 14
US-09-918-995-22675
; Sequence 22675, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22675
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-22675

  Query Match             29.3%;  Score 369.8;  DB 10;  Length 488;
  Best Local Similarity   99.5%;  Pred. No. 3.7e-67;
  Matches  371;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         35 AGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGG 94
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         28 AGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGAGCGTCAGAGACACGG 87

Qy         95 ACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCAT 154
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         88 ACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCAT 147

Qy        155 TTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTG 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 TTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTG 207

Qy        215 TCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTG 274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 TCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTG 267

Qy        275 ACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAA 334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 ACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAA 327

Qy        335 AAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTG 394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 AAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTG 387

Qy        395 TCATGATGAAGAT 407
              |||||||||||
Db        388 TCATGATGAAGCT 400


RESULT 15
US-09-918-995-10303
; Sequence 10303, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10303
;   LENGTH: 458
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(458)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10303

  Query Match             26.6%;  Score 336;  DB 10;  Length 458;
  Best Local Similarity   91.1%;  Pred. No. 4e-60;
  Matches  368;  Conservative    0;  Mismatches   35;  Indels    1;  Gaps    1;

Qy        639 GAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCG 698
              ||||||   | || ||||||||||||||| ||||||||||||||||| |||| ||||
Db         56 GAGAAAGACTGACAATATATTTACTGAGACGTTAAAGGAAGACAGTATTACTGCACCTCT 115

Qy        699 CTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGA 758
              | ||   ||  ||     | ||||   |||  ||||| ||||| | |||  |||| |||
Db        116 GTGTGCATGCGCTA-CCCCACCAACCCGACCAGTCGAGAATTCCAATCAGTGCTTCCTGA 174

Qy        759 ATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAAT 818
              ||  |||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Db        175 ATTGCAAGTAACACAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAAT 234

Qy        819 GAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTT 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 GAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTT 294

Qy        879 GAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATA 938
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 GAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATA 354

Qy        939 TAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTG 998
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 TAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTG 414
```

```
Qy        999 CCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTT 1042
              ||||||||||||||||||||||||||||||||||||||||||||
Db        415 CCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTT 458
```

Search completed: July 8, 2005, 07:35:57
Job time : 1946 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 09:23:13 ; Search time 861 Seconds
(without alignments)
9209.653 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 6330943 seqs, 3139157217 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1263	100.0	1263	17	US-10-364-505-1	Sequence 1, Appli
	2	1263	100.0	1263	19	US-10-681-199-1	Sequence 1, Appli
	3	1263	100.0	1993	17	US-10-364-505-2	Sequence 2, Appli
	4	1263	100.0	1993	19	US-10-681-199-2	Sequence 2, Appli
	5	1002	79.3	1263	19	US-10-681-199-19	Sequence 19, Appl
	6	998	79.0	1559	17	US-10-108-260A-575	Sequence 575, App
	7	960	76.0	1263	19	US-10-681-199-13	Sequence 13, Appl
	8	859	68.0	1263	19	US-10-681-199-15	Sequence 15, Appl
	9	696	55.1	1263	19	US-10-681-199-17	Sequence 17, Appl
	10	326	25.8	488	10	US-09-918-995-22675	Sequence 22675, A
c	11	325	25.7	325	18	US-10-240-425-872	Sequence 872, App
	12	270	21.4	458	10	US-09-918-995-10303	Sequence 10303, A
c	13	253	20.0	715	21	US-10-956-157-3114	Sequence 3114, Ap
	14	253	20.0	715	21	US-10-956-157-8349	Sequence 8349, Ap
	15	214	16.9	464	17	US-10-242-535A-30001	Sequence 30001, A
	16	214	16.9	464	18	US-10-085-783A-30001	Sequence 30001, A
	17	184	14.6	50000	17	US-10-364-505-7	Sequence 7, Appli
	18	184	14.6	50000	19	US-10-681-199-7	Sequence 7, Appli
	19	155	12.3	50000	17	US-10-364-505-8	Sequence 8, Appli
	20	155	12.3	50000	19	US-10-681-199-8	Sequence 8, Appli
	21	100	7.9	505	9	US-09-783-590-8049	Sequence 8049, Ap
	22	33	2.6	1697	17	US-10-364-505-4	Sequence 4, Appli
	23	33	2.6	1697	19	US-10-681-199-4	Sequence 4, Appli
c	24	26	2.1	708	13	US-10-027-632-24581	Sequence 24581, A
c	25	26	2.1	708	13	US-10-027-632-24582	Sequence 24582, A
c	26	26	2.1	708	17	US-10-027-632-24581	Sequence 24581, A
c	27	26	2.1	708	17	US-10-027-632-24582	Sequence 24582, A
c	28	25	2.0	25	17	US-10-364-505-11	Sequence 11, Appl
	29	25	2.0	25	17	US-10-364-505-12	Sequence 12, Appl
c	30	25	2.0	25	19	US-10-681-199-11	Sequence 11, Appl
	31	25	2.0	25	19	US-10-681-199-12	Sequence 12, Appl
	32	25	2.0	25	21	US-10-956-157-79342	Sequence 79342, A
	33	25	2.0	25	21	US-10-956-157-79343	Sequence 79343, A
	34	25	2.0	25	21	US-10-956-157-79344	Sequence 79344, A
	35	25	2.0	25	21	US-10-956-157-79348	Sequence 79348, A
	36	25	2.0	25	21	US-10-956-157-79350	Sequence 79350, A
	37	25	2.0	25	21	US-10-956-157-79356	Sequence 79356, A
	38	25	2.0	25	21	US-10-956-157-79357	Sequence 79357, A
	39	25	2.0	25	21	US-10-956-157-79359	Sequence 79359, A
	40	25	2.0	25	21	US-10-956-157-79360	Sequence 79360, A
	41	25	2.0	25	21	US-10-956-157-79361	Sequence 79361, A
	42	25	2.0	25	21	US-10-956-157-79362	Sequence 79362, A
c	43	22	1.7	462	20	US-10-425-115-115886	Sequence 115886,
	44	22	1.7	589	13	US-10-027-632-214594	Sequence 214594,
	45	22	1.7	589	17	US-10-027-632-214594	Sequence 214594,

ALIGNMENTS

RESULT 1
US-10-364-505-1
; Sequence 1, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
; APPLICANT: Kere, Juha
; APPLICANT: Taipale, Mikko
; APPLICANT: No. US20030219787Alola-Hemmi, Jaana
; APPLICANT: Kaminen, Nina
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0199P
; CURRENT APPLICATION NUMBER: US/10/364,505
; CURRENT FILING DATE: 2003-02-12
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequence for human DYXC1 (cDNA)
US-10-364-505-1

Query Match 100.0%; Score 1263; DB 17; Length 1263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
```

```
Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263
```

RESULT 2
US-10-681-199-1
; Sequence 1, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: coding sequence for human DYXC1 (cDNA)
US-10-681-199-1

Query Match 100.0%; Score 1263; DB 19; Length 1263;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 540
```

```
Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260

Qy       1261 TAA 1263
              |||
Db       1261 TAA 1263


RESULT 3
US-10-364-505-2
; Sequence 2, Application US/10364505
; Publication No. US20030219787A1
; GENERAL INFORMATION:
;  APPLICANT: Kere, Juha
;  APPLICANT:  Taipale, Mikko
;  APPLICANT:  No. US20030219787A1ola-Hemmi, Jaana
;  APPLICANT:  Kaminen, Nina
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0199P
;  CURRENT APPLICATION NUMBER: US/10/364,505
;  CURRENT FILING DATE:  2003-02-12
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1993
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (369)..(1628)
;   FEATURE:
;   OTHER INFORMATION: human DYXC1 mRNA as cDNA
US-10-364-505-2

  Query Match            100.0%;  Score 1263;  DB 17;  Length 1993;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328
```

```
Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631


RESULT 4
US-10-681-199-2
; Sequence 2, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 1993
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (369)..(1628)
;   FEATURE:
;   OTHER INFORMATION: human DYXC1 mRNA as cDNA
US-10-681-199-2

  Query Match             100.0%;  Score 1263;  DB 19;  Length 1993;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1263;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 428

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 488

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 548

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 608

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 668

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 728

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 788

Qy        421 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 AGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 848

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAA 908

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTT 968

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 1028

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1029 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 1088

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1089 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 1148

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1149 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 1208

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1209 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 1268

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1269 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1328

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1388

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1389 CACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGAC 1448

Qy       1081 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1449 AATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAA 1508

Qy       1141 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1509 TTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAA 1568

Qy       1201 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1569 ATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCT 1628

Qy       1261 TAA 1263
              |||
Db       1629 TAA 1631


RESULT 5
US-10-681-199-19
; Sequence 19, Application US/10681199
; Publication No. US20040138441A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
;   LENGTH: 1263
;   TYPE: DNA
;   ORGANISM: Pan paniscus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1260)
US-10-681-199-19

  Query Match             79.3%;  Score 1002;  DB 19;  Length 1263;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1252;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66

Qy         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126

Qy        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186

Qy        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246

Qy        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAA 306

Qy        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366

Qy        367 CGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        367 CGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426

Qy        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGG 486
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAAGCCTGG 486

Qy        487 AAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGT 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 AAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGT 546

Qy        547 CAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGA 606
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        547 CAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGA 606

Qy        607 AATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAG 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 AATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAG 666

Qy        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726

Qy        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786

Qy        787 CTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CTGCACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846

Qy        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906

Qy        907 ACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966

Qy        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026

Qy       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086

Qy       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146

Qy       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206

Qy       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263


RESULT 6
US-10-108-260A-575
; Sequence 575, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 575
;   LENGTH: 1559
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-108-260A-575

  Query Match             79.0%;  Score 998;  DB 17;  Length 1559;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1048;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 102

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        103 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 162

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        163 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 222

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 282

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        283 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 342

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 402

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        403 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 462

Qy        421 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 AGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAA 522

Qy        481 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 GCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAA 582

Qy        541 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTT 600
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        583 TTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTT 642

Qy        601 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        643 ACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTT 702

Qy        661 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        703 ACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATC 762

Qy        721 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        763 AACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAG 822

Qy        781 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 GAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTT 882

Qy        841 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 TGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTG 942

Qy        901 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 TTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAAT 1002

Qy        961 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 AATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTA 1062

Qy       1021 CACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              |||||||||||||||||||||||||||||
Db       1063 CACAAGGCTATTGAAGATTCTTCTAAGGC 1091
```

RESULT 7
US-10-681-199-13
; Sequence 13, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
; APPLICANT: KERE, Juha
; TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
; FILE REFERENCE: 0933-0214P
; CURRENT APPLICATION NUMBER: US/10/681,199
; CURRENT FILING DATE: 2003-10-09
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Pan troglodytes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1260)
US-10-681-199-13

Query Match 76.0%; Score 960; DB 19; Length 1263;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1210; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         49 GTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACG 108
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db         49 GTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCATG 108

Qy        109 GAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 GAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 168

Qy        169 ATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTAT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTAT 228

Qy        229 AAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 AAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288

Qy        289 CAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 CAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAA 348

Qy        349 GCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATT 408
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        349 GCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATT 408

Qy        409 GAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACT 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACT 468

Qy        469 AAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAG 528
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 AAAGAATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAG 528

Qy        529 AGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAAA 588
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        529 AGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAAATAAAA 588

Qy        589 TATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCA 648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        589 TATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCA 648

Qy        649 GAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGC 708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        649 GAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGC 708

Qy        709 AGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTA 768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        709 AGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTA 768

Qy        769 GCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGAC 828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        769 GCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGAC 828

Qy        829 ATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAA 888
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        829 ATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAA 888

Qy        889 GGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCC 948
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        889 GGAAACAAATTGTTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCC 948

Qy        949 ATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAA 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        949 ATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAA 1008

Qy       1009 CTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCA 1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1009 CTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCA 1068

Qy       1069 CCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTC 1128
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1069 CCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTC 1128

Qy       1129 TGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGAT 1188
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1129 TGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGAT 1188

Qy       1189 CCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACA 1248
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1189 CCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACA 1248

Qy       1249 GAACTAAAATCTTAA 1263
             |||||||||||||||
Db       1249 GAACTAAAATCTTAA 1263


RESULT 8
US-10-681-199-15
; Sequence 15, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 1263
;   TYPE: DNA
;   ORGANISM: Gorilla gorilla
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1260)
US-10-681-199-15

  Query Match             68.0%;  Score 859;  DB 19;  Length 1263;
  Best Local Similarity   99.4%;  Pred. No. 0;
  Matches 1209;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         48 GGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCAC 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         48 GGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCAC 107

Qy        108 GGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCC 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 GGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCC 167

Qy        168 CATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTA 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTA 227

Qy        228 TAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGAT 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        228 TAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGAT 287

Qy        288 GCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGA 347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 GCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGA 347

Qy        348 AGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGAT 407
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        348 AGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGAT 407

Qy        408 TGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCAC 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        408 TGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCAC 467

Qy        468 TAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCA 527
              ||||| |||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        468 TAAAGAATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAGAAATTCA 527

Qy        528 GAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAATAAA 587
              |||||||||||| |||||||||||||||||||||||||||||||| |||||||||| ||||
Db        528 GAGAGAAGAGAAGTTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAATTAAA 587

Qy        588 ATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTC 647
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        588 ATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTC 647

Qy        648 AGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGG 707
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        648 AGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGG 707

Qy        708 CAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGT 767
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        708 CAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGT 767

Qy        768 AGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGA 827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 AGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGA 827

Qy        828 CATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAA 887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        828 CATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAA 887

Qy        888 AGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGC 947
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        888 AGGAAACAAATTGTTTGCAACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGC 947

Qy        948 CATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAA 1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        948 CATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAA 1007

Qy       1008 ACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCC 1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1008 ACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCC 1067

Qy       1068 ACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATT 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1068 ACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATT 1127

Qy       1128 CTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGA 1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 CTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGA 1187

Qy       1188 TCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAAC 1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 TCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAAC 1247

Qy       1248 AGAACTAAAATCTTAA 1263
              ||||||||||||||||
Db       1248 AGAACTAAAATCTTAA 1263


RESULT 9
US-10-681-199-17
; Sequence 17, Application US/10681199
; Publication No. US20040138441A1
; GENERAL INFORMATION:
;  APPLICANT: KERE, Juha
;  TITLE OF INVENTION: NOVEL HUMAN GENE FUNCTIONALLY RELATED TO DYSLEXIA
;  FILE REFERENCE: 0933-0214P
;  CURRENT APPLICATION NUMBER: US/10/681,199
;  CURRENT FILING DATE:  2003-10-09
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 1263
;   TYPE: DNA
;   ORGANISM: Pongo pygmaeus
;   FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (1)..(1260)
US-10-681-199-17

  Query Match             55.1%;  Score 696;  DB 19;  Length 1263;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 1246;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCC 66

Qy         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTC 126

Qy        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 AACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGC 186

Qy        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATG 246

Qy        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAA 306
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        247 TGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGACGATGCAAAGAATTAGAGAAAAA 306

Qy        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 TCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAG 366

Qy        367 CGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        367 CGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAA 426

Qy        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGG 486
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        427 AAAATAGAAGATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGAATTGGAAGCCTGG 486

Qy        487 AAAGAATATCAAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGT 546
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        487 AAAGAATATCAAAGAAAAGCTGAGGAGCACAAAAAAATTCAGAGAGAAGAGAAATTATGT 546

Qy        547 CAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGA 606
              |||||||||||||||||||||||| |||||||||| ||||||| ||||||||||||||||
Db        547 CAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAAAAATTAAAATACAAGAGTCTTACTAGA 606

Qy        607 AATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAG 666
              |||| |||||||||||| |||||||||||||| |||||||||||||||||||||||||||
Db        607 AATTCGGCATCTAGAAACCTTGCTCCAAAAGGAAGAAATTCAGAAAATATATTTACTGAG 666

Qy        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 AAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTT 726

Qy        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 ACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGG 786

Qy        787 CTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        787 CTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGAT 846

Qy        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        847 TTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCA 906

Qy        907 ACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        907 ACAGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAG 966

Qy        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        967 ATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAG 1026

Qy       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1027 GCTATTGAAGATTCTTCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCT 1086

Qy       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1087 AATGCAAGAATGAAGGCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTAT 1146

Qy       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1147 GTAGAAGGCCTACAGGATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTA 1206

Qy       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1207 CAAATTGATGCTGAGAAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263


RESULT 10
US-09-918-995-22675
; Sequence 22675, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 22675
;   LENGTH: 488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-918-995-22675

  Query Match             25.8%;  Score 326;  DB 10;  Length 488;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-151;
  Matches  326;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 GCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCAT 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 GCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCAT 132

Qy        140 TTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTG 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        133 TTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTG 192

Qy        200 GGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTT 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 GGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTT 252

Qy        260 CTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAG 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 CTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAG 312

Qy        320 CACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAA 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAA 372

Qy        380 AATACGCACTAAGTGTCATGATGAAG 405
              |||||||||||||||||||||||||
Db        373 AATACGCACTAAGTGTCATGATGAAG 398


RESULT 11
US-10-240-425-872/c
; Sequence 872, Application US/10240425
```

```
; Publication No. US20040033502A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Amanda
;  APPLICANT:  Boland, Joseph F.
;  APPLICANT:  Lord, Reginald V.
;  APPLICANT:  Alvarez, Chris
;  APPLICANT:  Wetzel, Jon C.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Vockley, Joseph G.
;  TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
;  FILE REFERENCE: 44921-5026
;  CURRENT APPLICATION NUMBER: US/10/240,425
;  CURRENT FILING DATE:  2002-09-30
;  PRIOR APPLICATION NUMBER: PCT/US01/09847
;  PRIOR FILING DATE: 2001-03-28
;  PRIOR APPLICATION NUMBER: US 60/193,446
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 1588
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 872
;   LENGTH: 325
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20040033502A1 AI783611
US-10-240-425-872

  Query Match             25.7%;  Score 325;  DB 18;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-151;
  Matches  325;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAA 266

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 206

Qy        644 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 146

Qy        704 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 86

Qy        764 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 26

Qy        824 CTGACATAGCTGAACTTTGCGATTT 848
              |||||||||||||||||||||||||
Db         25 CTGACATAGCTGAACTTTGCGATTT 1


RESULT 12
US-09-918-995-10303
; Sequence 10303, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10303
;   LENGTH: 458
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(458)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-10303

  Query Match             21.4%;  Score 270;  DB 10;  Length 458;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-123;
  Matches  270;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        773 AAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAG 832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 AAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATACTGACATAG 248

Qy        833 CTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAA 892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 CTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTGAAGGATAAAGGAA 308

Qy        893 ACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAA 952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 ACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATATAATTTAGCCATAA 368

Qy        953 GACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAA 1012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 GACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGCCACCTAAAACTAA 428

Qy       1013 AAAACTTACACAAGGCTATTGAAGATTCTT 1042
              |||||||||||||||||||||||||||||
Db        429 AAAACTTACACAAGGCTATTGAAGATTCTT 458


RESULT 13
US-10-956-157-3114/c
; Sequence 3114, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3114
;   LENGTH: 715
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-3114

  Query Match             20.0%;  Score 253;  DB 21;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-115;
  Matches  253;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        638 GGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTC 697
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTC 371

Qy        698 GCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTG 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 GCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTG 311

Qy        758 AATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAA 817
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 AATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAA 251

Qy        818 TGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGT 877
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 TGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGT 191
```

```
Qy        878 TGAAGGATAAAGG 890
              |||||||||||||
Db        190 TGAAGGATAAAGG 178


RESULT 14
US-10-956-157-8349
; Sequence 8349, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
;  APPLICANT: Wyeth
;  APPLICANT:  Mounts, William
;  TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
;  TITLE OF INVENTION:  HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
;  FILE REFERENCE: 031896-043000 (AM 101081)
;  CURRENT APPLICATION NUMBER: US/10/956,157
;  CURRENT FILING DATE:  2004-10-04
;  NUMBER OF SEQ ID NOS: 319805
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8349
;   LENGTH: 715
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-956-157-8349

  Query Match             20.0%;  Score 253;  DB 21;  Length 715;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-115;
  Matches  253;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        638 GGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTC 697
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTC 345

Qy        698 GCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTG 757
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 GCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTG 405

Qy        758 AATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAA 817
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAA 465

Qy        818 TGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGT 877
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 TGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGT 525

Qy        878 TGAAGGATAAAGG 890
              |||||||||||||
Db        526 TGAAGGATAAAGG 538


RESULT 15
US-10-242-535A-30001
; Sequence 30001, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30001
;   LENGTH: 464
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (263)..(263)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (321)..(321)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (347)..(347)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-30001

  Query Match             16.9%;  Score 214;  DB 17;  Length 464;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-95;
  Matches  214;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 108

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        109 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 168

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 228

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTG 214
              |||||||||||||||||||||||||||||||||
Db        229 AGCAGCAAAGCAAAGATTGGGAATGACACCATTG 262


Search completed: July  8, 2005, 16:12:46
Job time : 865 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 06:20:46 ; Search time 9619 Seconds
(without alignments)
4997.939 Million cell updates/sec

Title: US-10-681-199-1
Perfect score: 1263
Sequence: 1 atgcctcttcaggttagcga.....gaacagaactaaaatcttaa 1263

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	838	66.3	1600	3	BC017392	BC017392 Homo sapi
	2	583	46.2	844	6	CD107587	CD107587 AGENCOURT
	3	462	36.6	535	7	CN478982	CN478982 UI-CF-FN0
	4	446	35.3	1131	5	BQ217312	BQ217312 AGENCOURT
	5	422	33.4	462	7	CN429788	CN429788 170006000
	6	410	32.5	788	2	BE564350	BE564350 601343161
	7	405	32.1	823	5	BU567908	BU567908 AGENCOURT
	8	378	29.9	792	4	BG771796	BG771796 602720472
	9	374	29.6	933	6	CD358543	CD358543 AGENCOURT
	10	364	28.8	573	2	BE972748	BE972748 601652170
c	11	342	27.1	772	6	CA422987	CA422987 UI-H-FL0-
	12	340	26.9	468	6	CB297042	CB297042 12B22058_
	13	338	26.8	417	1	AI024221	AI024221 ov81e02.s
c	14	325	25.7	325	1	AI783611	AI783611 tz99c01.x
	15	323	25.6	651	5	BU852580	BU852580 AGENCOURT
	16	320	25.3	701	4	BG482634	BG482634 602502783
	17	313	24.8	1183	2	BF207765	BF207765 601861861
	18	290	23.0	665	4	BG540324	BG540324 602568825
	19	275	21.8	559	5	BU607404	BU607404 UI-CF-FN0
	20	274	21.7	545	2	BF216970	BF216970 601884034
c	21	251	19.9	625	5	BM984145	BM984145 UI-CF-DU1
	22	250	19.8	270	2	BF372375	BF372375 PM4-FT002
	23	244	19.3	744	4	BG192162	BG192162 RST11269
c	24	241	19.1	267	1	AI073572	AI073572 ov45g11.x
c	25	218	17.3	583	6	CB160627	CB160627 K-EST0220
c	26	216	17.1	382	1	AA724419	AA724419 ah91f07.s
	27	215	17.0	775	4	BG206612	BG206612 RST26063
c	28	212	16.8	523	2	BE178005	BE178005 RC3-HT060
	29	197	15.6	251	1	AA317003	AA317003 EST18868
c	30	182	14.4	454	2	BE463906	BE463906 hy18f08.x
	31	175	13.9	988	4	BI517373	BI517373 603041624
	32	150	11.9	752	4	BG183140	BG183140 RST2158 A
c	33	148	11.7	940	4	BG220651	BG220651 RST40439
	34	146	11.6	855	2	BF248143	BF248143 601859338
c	35	105	8.3	467	1	AI674107	AI674107 wd18c04.x
c	36	102	8.1	500	4	BI517981	BI517981 603041624
c	37	85	6.7	459	5	BX095509	BX095509 BX095509
	38	78	6.2	402	2	BE178258	BE178258 RC3-HT060
	39	78	6.2	683	5	BP460416	BP460416 BP460416
	40	70	5.5	112	4	BM820152	BM820152 K-EST0088
c	41	64	5.1	440	1	AI360851	AI360851 qx99h12.x
	42	61	4.8	793	7	CO738958	CO738958 S1LE04c21
c	43	60	4.8	654	5	BM971229	BM971229 UI-CF-DU1
c	44	57	4.5	515	5	BU738973	BU738973 UI-E-EJ0-
	45	51	4.0	285	8	AQ067717	AQ067717 HS_2239_B

ALIGNMENTS

RESULT 1
BC017392
LOCUS BC017392 1600 bp mRNA linear HTC 23-AUG-2002
DEFINITION Homo sapiens, Similar to RIKEN cDNA 1700010I24 gene, clone IMAGE:4081622, mRNA.
ACCESSION BC017392
VERSION BC017392.1 GI:19263480
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 32 Row: k Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction
This clone has the following problem: frame shifted.
FEATURES Location/Qualifiers
source 1. .1600
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4081622"
/tissue_type="Bladder, carcinoma"
/clone_lib="NIH_MGC_53"
/lab_host="DH10B"

```
                     /note="Vector: pDNR-LIB"
ORIGIN

  Query Match            66.3%;  Score 838;  DB 3;  Length 1600;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1008;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         41 AGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGT 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         93 AGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGT 152

Qy        101 TCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTT 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        153 TCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTT 212

Qy        161 ATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCA 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 ATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCA 272

Qy        221 CCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAG 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 CCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAG 332

Qy        281 AGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAG 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAG 392

Qy        341 CTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGA 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 CTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGA 452

Qy        401 TGAAGATTGAAGAAGAAGAGAGG-AAAAAAATAGAAGATATGAAAGAAAATGAACGGATA 459
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        453 TGAAGATTGAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAAATGAACGGATA 512

Qy        460 AAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAA 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 AAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAGGAGCAAAAA 572

Qy        520 AAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAAGCAAATTAAAGAAGGAAGAAAA 579
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        573 AAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGAAAGAAAA 632

Qy        580 AAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGG 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGG 692

Qy        640 AGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGC 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 AGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGC 752

Qy        700 TCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAA 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAA 812

Qy        760 TCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATG 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 TCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATG 872

Qy        820 AATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTG 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 AATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGAACCCAGAATGGTTG 932

Qy        880 AAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATAT 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 AAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATGCATAT 992

Qy        940 AATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGC 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        993 AATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTGCTTGC 1052

Qy       1000 CACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGC 1049
              ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1053 CACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGC 1102


RESULT 2
CD107587
LOCUS       CD107587                844 bp    mRNA    linear   EST 15-MAY-2003
DEFINITION  AGENCOURT_14021179 NIH_MGC_179 Homo sapiens cDNA clone
            IMAGE:30368368 5', mRNA sequence.
ACCESSION   CD107587
VERSION     CD107587.1  GI:30760665
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 844)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM431  row: a  column: 17
            High quality sequence stop: 634.
FEATURES             Location/Qualifiers
     source          1. .844
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30368368"
                     /tissue_type="Pituitary"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_179"
                     /note="Organ: brain; Vector: pCMV-SPORT6.1; Site_1: EcoRV
                     (destroyed); Site_2: NotI; Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning).  Average insert size 1.1 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match            46.2%;  Score 583;  DB 6;  Length 844;
  Best Local Similarity  99.7%;  Pred. No. 5.1e-287;
  Matches  753;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         18 CGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGT 77
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGT 60

Qy         78 GTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCC 137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCC 120

Qy        138 ATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGAT 197
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGAT 180

Qy        198 TGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCT 257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCT 240

Qy        258 TTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACA 317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACA 300

Qy        318 AGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCA 377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 AGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCA 360

Qy        378 AAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGA 437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGA 420

Qy        438 TATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCA 497
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCA 480

Qy        498 AAGAAAAGCTGAGGAGC-AAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAA 556
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        481 AAGAAAAGCTGAGGAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAA 540

Qy        557 AGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCAT 616
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCAT 600

Qy        617 CTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGG 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGG 660

Qy        677 AAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAG 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAG 720

Qy        737 TATTCCCAACAGCTCTTCGTGAATCACAAGTAGCA 771
              |||||||||||||||||||||||||||||||||||
Db        721 TATTCCCAACAGCTCTTCGTGAATCACAAGTAGCA 755
```

RESULT 3
CN478982
LOCUS CN478982 535 bp mRNA linear EST 23-APR-2004
DEFINITION UI-CF-FN0-afu-1-16-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone UI-CF-FN0-afu-1-16-0-UI 3', mRNA sequence.
ACCESSION CN478982
VERSION CN478982.1 GI:46547981
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/cf.html
The following repetitive elements were found in this cDNA sequence: 133-243, >(GAAAA)n#Simple_repeat 493-514, >AT_rich#Low_complexity (matched compliment)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FN0-afu-1-16-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FN0"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-FN0 is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-FN0
TAG_SEQ=GGCTGTAGGC"
ORIGIN

Query Match 36.6%; Score 462; DB 7; Length 535;
Best Local Similarity 100.0%; Pred. No. 5.4e-225;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        391 AGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAAT 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         39 AGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAAT 98

Qy        451 GAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAG 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTGAG 158

Qy        511 GAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAA 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 GAGCAAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAA 218

Qy        571 GGAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCT 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 GGAAGAAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCT 278

Qy        631 CCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCT 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 CCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCT 338

Qy        691 GCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCT 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 GCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCT 398

Qy        751 CTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGA 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 CTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGA 458

Qy        811 AGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAA 852
              |||||||||||||||||||||||||||||||||||||||||
Db        459 AGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAA 500
```

RESULT 4
BQ217312
LOCUS BQ217312 1131 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7558431 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6047052 5', mRNA sequence.
ACCESSION BQ217312
VERSION BQ217312.1 GI:20398712
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1131)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP

```
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM13293  row: n  column: 13
           High quality sequence stop: 388.
FEATURES             Location/Qualifiers
     source          1. .1131
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6047052"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_72"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
ORIGIN

  Query Match             35.3%;  Score 446;  DB 5;  Length 1131;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-217;
  Matches  446;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         42 GACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTT 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GACTGCGGTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTT 60

Qy        102 CTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTA 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCACGGAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTA 120

Qy        162 TGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCAC 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCAC 180

Qy        222 CTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGA 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGA 240

Qy        282 GATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGC 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGC 300

Qy        342 TACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGAT 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGAT 360

Qy        402 GAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAA 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGATTGAAGAAGAAGAGAGGAAAAAAATAGAAGATATGAAAGAAAATGAACGGATAAA 420

Qy        462 AGCCACTAAAGCATTGGAAGCCTGGA 487
              |||||||||||||||||||||||||
Db        421 AGCCACTAAAGCATTGGAAGCCTGGA 446


RESULT 5
CN429788
LOCUS       CN429788                462 bp    mRNA    linear   EST 16-MAY-2004
DEFINITION  17000600026401 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN429788
VERSION     CN429788.1  GI:47417382
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 462)
  AUTHORS   Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
            Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
            Lebkowski,J and Stanton,L.W.
  TITLE     Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
  JOURNAL   Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 462   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .462
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, DMSO-treated H9 cell
                     line"
                     /clone_lib="GRN_PREHEP"
                     /note="oligo dT primed, full-length enriched cDNA library
                     from DMSO-treated hES cell line H9 (p22) maintained in
                     feeder-free conditions"
ORIGIN

  Query Match             33.4%;  Score 422;  DB 7;  Length 462;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-204;
  Matches  422;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 100

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 160

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 220

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 280

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 340

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 400

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 460

Qy        421 AG 422
              ||
Db        461 AG 462


RESULT 6
BE564350
LOCUS       BE564350                788 bp    mRNA    linear   EST 15-AUG-2000
DEFINITION  601343161F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3685335 5',
            mRNA sequence.
ACCESSION   BE564350
VERSION     BE564350.1  GI:9808070
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 788)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM373  row: i  column: 16
            High quality sequence stop: 556.
FEATURES             Location/Qualifiers
     source          1. .788
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3685335"
                     /tissue_type="carcinoma, cell line"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_53"
                     /note="Organ: bladder; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI
                     (ggccattatggcc); Double-stranded cDNA was prepared from
                     cell line RNA.  5' and 3' adaptors were used in cloning as
                     follows: 5' adaptor sequence:  5'-CACGGCCATTATGGCC-3' and
                     3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T).  Average insert size 1.55
                     kb (range 0.9-4.0 kb).  15/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA)."
ORIGIN

  Query Match             32.5%;  Score 410;  DB 2;  Length 788;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-198;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         23 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 82

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 142

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 202

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 262

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 322

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 382

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGA 410
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGA 432


RESULT 7
BU567908
LOCUS       BU567908                823 bp    mRNA    linear   EST 16-SEP-2002
DEFINITION  AGENCOURT_10398893 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6614647
            5', mRNA sequence.
ACCESSION   BU567908
VERSION     BU567908.1  GI:22918208
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 823)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2855  row: h  column: 07
            High quality sequence stop: 540.
FEATURES             Location/Qualifiers
     source          1. .823
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6614647"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match             32.1%;  Score 405;  DB 5;  Length 823;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-196;
  Matches  405;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 111

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 171

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 231

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 291

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 351

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 411
```

```
Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAG 405
              ||||||||||||||||||||||||||||||||||||||||||||
Db        412 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAG 456


RESULT 8
BG771796
LOCUS       BG771796                792 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602720472F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4837505 5',
            mRNA sequence.
ACCESSION   BG771796
VERSION     BG771796.1  GI:14082449
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 792)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
             cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10769  row: p  column: 18
            High quality sequence stop: 671.
FEATURES             Location/Qualifiers
     source          1. .792
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4837505"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_97"
                     /note="Organ: testis; Vector: pBluescriptR (modified
                     pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
                     (gtcgag); Oligo-dT primed using primer
                     5'-TTTTTTTTTTTTTTTVN-3', size-selected for average
                     insert size 2.2 kb and normalized to ROT 5. This is a
                     primary library enriched for full-length clones and
                     constructed using the Cap-trapper method (Carninci, in
                     preparation). Library constructed by M. Brownstein
                     (NIMH/NHGRI, National Institutes of Health). Note: this is
                     a NIH_MGC Library."
ORIGIN

  Query Match            29.9%;  Score 378;  DB 4;  Length 792;
  Best Local Similarity  99.8%;  Pred. No. 6.4e-182;
  Matches  428;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         37 ATGCCTCTTCAGGTTAGCGATTACAGCTGGCAGCAGACGAAGACTGCGGTCTTTCTGTCT 96

Qy         61 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 CTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACGGAAAACTATCTG 156

Qy        121 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 AAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCCATAGACGATGAG 216

Qy        181 AGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTATAAAAAAGAAGCG 240
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        217 AGCAGCAAAGCAAAGATTGGGAATGACACCATGGTCTTCACCTTGTATAAAAAAGAAGCG 276

Qy        241 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATGCAAAGAATTAGA 336

Qy        301 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 GAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAAGCAAAAGCTGCA 396

Qy        361 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATTGAAGAAGAAGAG 456

Qy        421 AGGAAAAAA 429
              |||||||||
Db        457 AGGAAAAAA 465


RESULT 9
CD358543
LOCUS       CD358543                933 bp    mRNA    linear   EST 29-MAY-2003
DEFINITION  AGENCOURT_14255831 NIH_MGC_180 Homo sapiens cDNA clone
            IMAGE:30386203 5', mRNA sequence.
ACCESSION   CD358543
VERSION     CD358543.1  GI:31129978
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM454  row: h  column: 20
            High quality sequence start: 303
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30386203"
                     /lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
                     /clone_lib="NIH_MGC_180"
                     /note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI;
                     Site_2: EcoRV (destroyed); Library is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed upon
                     cloning). Average insert size 1.68 kb. Library was
                     constructed by (Invitrogen). Note: this is a NIH_MGC
                     Library."
ORIGIN

  Query Match            29.6%;  Score 374;  DB 6;  Length 933;
  Best Local Similarity  99.8%;  Pred. No. 7.3e-180;
  Matches  424;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        437 ATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATC 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 ATATGAAAGAAAATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATC 372
```

```
Qy        497 AAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAA 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 AAAGAAAAGCTGAGGAGCAAAAAAAAATTCAGAGAGAAGAGAAATTATGTCAAAAAGAAA 432

Qy        557 AGCAAATTAAAGAAGGAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCAT 616
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        433 AGCAAATTAAAGAAGAAAGAAAAAAAATAAAATATAAGAGTCTTACTAGAAATTTGGCAT 492

Qy        617 CTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGG 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 CTAGAAATCTTGCTCCAAAAGGGAGAAATTCAGAAAATATATTTACTGAGAAGTTAAAGG 552

Qy        677 AAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAG 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 AAGACAGTATTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAG 612

Qy        737 TATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAAC 796
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 TATTCCCAACAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAAC 672

Qy        797 AAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAG 856
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        673 AAGCTGAGGCACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAG 732

Qy        857 AAGAA 861
              |||||
Db        733 AAGAA 737
```

RESULT 10
BE972748

102(b)

```
LOCUS       BE972748                573 bp    mRNA    linear   EST 04-OCT-2000
DEFINITION  601652170F1 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:3935392 5',
            mRNA sequence.
ACCESSION   BE972748
VERSION     BE972748.1  GI:10586084
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONETECH Laboratories, Inc.
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM777  row: 1  column: 17
            High quality sequence stop: 571.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3935392"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN
```

```
  Query Match             28.8%;  Score 364;  DB 2;  Length 573;
  Best Local Similarity   100.0%;  Pred. No. 9.4e-175;
  Matches  364;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        686 TTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAA 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCCTGCTCCTCGCTCTGTTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAA 60

Qy        746 CAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGG 805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGCTCTTCGTGAATCACAAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGG 120

Qy        806 CACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGA 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACGAAGAGCAATGAATACTGACATAGCTGAACTTTGCGATTTAAAAGAAGAAGAAAAGA 180

Qy        866 ACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAG 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCCAGAATGGTTGAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAG 240

Qy        926 CTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGA 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGA 300

Qy        986 ACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTA 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTA 360

Qy       1046 AGGC 1049
              ||||
Db        361 AGGC 364
```

```
RESULT 11
CA422987/c
LOCUS       CA422987                772 bp    mRNA    linear   EST 07-NOV-2002
DEFINITION  UI-H-FL0-bdk-g-13-0-UI.s1 NCI_CGAP_FL0 Homo sapiens cDNA clone
            UI-H-FL0-bdk-g-13-0-UI 3', mRNA sequence.
ACCESSION   CA422987
VERSION     CA422987.1  GI:24785713
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 772)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 1-44, >AT_rich#Low_complexity (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .772
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FL0-bdk-g-13-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
```

/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FL0"
/note="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI_CGAP_FL0 is a cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GAGGTCGGTG. The cell line was provided by Dr James Martin from University of Iowa. TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FL0
TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 27.1%; Score 342; DB 6; Length 772;
Best Local Similarity 100.0%; Pred. No. 1.8e-163;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        922 GCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTAT 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        720 GCAGCTATCAATGCATATAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTAT 661

Qy        982 TTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCT 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        660 TTGAACCGGGCTGCTTGCCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCT 601

Qy       1042 TCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCTAATGCAAGAATGAAG 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 TCTAAGGCACTGGAATTATTGATGCCACCTGTTACAGACAATGCTAATGCAAGAATGAAG 541

Qy       1102 GCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTATGTAGAAGGCCTACAG 1161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 GCACATGTACGACGTGGAACAGCATTCTGTCAACTAGAATTGTATGTAGAAGGCCTACAG 481

Qy       1162 GATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTACAAATTGATGCTGAG 1221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 GATTATGAAGCGGCACTTAAGATTGATCCATCCAACAAAATTGTACAAATTGATGCTGAG 421

Qy       1222 AAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 1263
              |||||||||||||||||||||||||||||||||||||||||||
Db        420 AAGATTCGGAATGTAATTCAAGGAACAGAACTAAAATCTTAA 379
```

RESULT 12
CB297042
LOCUS CB297042 468 bp mRNA linear EST 28-FEB-2003
DEFINITION 12B22058_rev_1_F09_r_075.ab1 Chimpanzee brain library Koos Pan troglodytes cDNA clone 12B22058_rev_1_F09_r_075.ab1 5', mRNA sequence.
ACCESSION CB297042
VERSION CB297042.1 GI:28622472
KEYWORDS EST.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 468)
AUTHORS Hellmann,I., Zollner,S., Enard,W., Ebersberger,I., Nickel,B. and Paabo,S.
TITLE Selection on human genes as revealed by comparisons to chimpanzee cDNA
JOURNAL Genome Res. (2003) In press
COMMENT Contact: Paabo S
Evolutionary Genetics
Max-Planck-Institute for evolutionary Anthropology
Deutscher Platz 6, 04103 Leipzig, Germany
Tel: +49-(0)-341-3550 500
Fax: +49-(0)-341-3550 555
Email: paabo@eva.mpg.de
Seq primer: M13 reverse.
FEATURES Location/Qualifiers
source 1. .468
/organism="Pan troglodytes"
/mol_type="mRNA"
/db_xref="taxon:9598"
/clone="12B22058_rev_1_F09_r_075.ab1"
/sex="male"
/tissue_type="brain, presumably cortex"
/dev_stage="adult"
/lab_host="Epicurian Coli (TM) XL-10-Gold"
/clone_lib="Chimpanzee brain library Koos"
/note="Vector: pUChi; Site_1: SfiI-A; Site_2: SfiI-B; The library was prepared using the SMART cDNA library construction Kit (Clontech), doing only primer extension, but not PCR amplification of the cDNA. The only deviation from the published protocoll was that we cloned the cDNA into a plasmid Vector."

ORIGIN

Query Match 26.9%; Score 340; DB 6; Length 468;
Best Local Similarity 99.7%; Pred. No. 1.9e-162;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         49 GTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACG 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 GTCTTTCTGTCTCTGCCCCTCAAAGGCGTGTGCGTCAGAGACACGGACGTGTTCTGCACG 137

Qy        109 GAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 GAAAACTATCTGAAGGTCAACTTTCCTCCATTTTTATTTGAGGCATTTCTTTATGCTCCC 197

Qy        169 ATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTAT 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACACCATTGTCTTCACCTTGTAT 257

Qy        229 AAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGGGTGTTGACAAAGAGATGATG 317

Qy        289 CAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAA 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGAGAGCAAAAGAAGCTACAGAA 377

Qy        349 GCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCACTAAGTGTCATGATGAAGATT 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 GCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATATGCACTAAGTGTCATGATGAAGATT 437

Qy        409 GAAGAAGAAGAGAGGAAAAAAATAGAAGATA 439
              |||||||||||||||||||||||||||||||
Db        438 GAAGAAGAAGAGAGGAAAAAAATAGAAGATA 468
```

RESULT 13
AI024221
LOCUS AI024221 417 bp mRNA linear EST 18-JUN-1998
DEFINITION ov81e02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1643738 3', mRNA sequence.
ACCESSION AI024221
VERSION AI024221.1 GI:3239265
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 417)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

```
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .417
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1643738"
                     /sex="male"
                     /lab_host="DH10B"
                     /clone_lib="Soares_testis_NHT"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech
                     Laboratories, Inc., and primed with  a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match           26.8%;  Score 338;  DB 1;  Length 417;
  Best Local Similarity   99.7%;  Pred. No. 2e-161;
  Matches  388;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        149 AGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACA 208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGGCATTTCTTTATGCTCCCATAGACGATGAGAGCAGCAAAGCAAAGATTGGGAATGACA 60

Qy        209 CCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGG 268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCATTGTCTTCACCTTGTATAAAAAAGAAGCGGCCATGTGGGAGACCCTTTCTGTGACGG 120

Qy        269 GTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGA 328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTGTTGACAAAGAGATGATGCAAAGAATTAGAGAAAAATCTATTTTACAAGCACAAGAGA 180

Qy        329 GAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGGGAAGATCAAAAATACGCAC 388
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        181 GAGCAAAAGAAGCTACAGAAGCAAAAGCTGCAGCAAAGCGCGAAGATCAAAAATACGCAC 240

Qy        389 TAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAA 448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAAGTGTCATGATGAAGATTGAAGAAGAAGAGAGGAAAAAAAATAGAAGATATGAAAGAAA 300

Qy        449 ATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTG 508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATGAACGGATAAAAGCCACTAAAGCATTGGAAGCCTGGAAAGAATATCAAAGAAAAGCTG 360

Qy        509 AGGAGCAAAAAAAAAATTCAGAGAGAAGAG 537
              |||||||||||||||||||||||||||||
Db        361 AGGAGCAAAAAAAAAATTCAGAGAGAAGAG 389


RESULT 14
AI783611/c
LOCUS       AI783611                325 bp    mRNA    linear   EST 16-DEC-1999
DEFINITION  tz99c01.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2296704 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AI783611
VERSION     AI783611.1  GI:5325420
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 325)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 437   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 325.
FEATURES             Location/Qualifiers
     source          1. .325
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2296704"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
ORIGIN

  Query Match           25.7%;  Score 325;  DB 1;  Length 325;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-155;
  Matches  325;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        524 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAA 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 TTCAGAGAGAAGAGAAATTATGTCAAAAAGAAAAGCAAATTAAAGAAGGAAGAAAAAAAAA 266

Qy        584 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TAAAATATAAGAGTCTTACTAGAAATTTGGCATCTAGAAATCTTGCTCCAAAAGGGAGAA 206

Qy        644 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 ATTCAGAAAATATATTTACTGAGAAGTTAAAGGAAGACAGTATTCCTGCTCCTCGCTCTG 146

Qy        704 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TTGGCAGTATTAAAATCAACTTTACCCCTCGAGTATTCCCAACAGCTCTTCGTGAATCAC 86

Qy        764 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 AAGTAGCAGAAGAGGAGGAGTGGCTACACAAACAAGCTGAGGCACGAAGAGCAATGAATA 26

Qy        824 CTGACATAGCTGAACTTTGCGATTT 848
              |||||||||||||||||||||||||
Db         25 CTGACATAGCTGAACTTTGCGATTT 1


RESULT 15
```

```
BU852580
LOCUS       BU852580                651 bp    mRNA    linear   EST 16-OCT-2002
DEFINITION  AGENCOURT_10501009 NIH_MGC_82 Homo sapiens cDNA clone IMAGE:6619012
            5', mRNA sequence.
ACCESSION   BU852580
VERSION     BU852580.1  GI:24037543
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2866  row: n  column: 04
            High quality sequence stop: 364.
FEATURES             Location/Qualifiers
     source          1. .651
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6619012"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_82"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).   Average insert size
                     1.35 kb (range 0.9-4.0 kb).  14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech  Laboratories (Palo
                     Alto, CA)."
ORIGIN

  Query Match            25.6%;  Score 323;  DB 5;  Length 651;
  Best Local Similarity   99.7%;  Pred. No. 1e-153;
  Matches  373;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        875 GGTTGAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATG 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGTTGAAGGATAAAGGAAACAAATTGTTTGCAACGGAAAACTATTTGGCAGCTATCAATG 62

Qy        935 CATATAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTG 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CATATAATTTAGCCATAAGACTAAATAATAAGATGCCACTATTGTATTTGAACCGGGCTG 122

Qy        995 CTTGCCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGG 1054
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CTCGCCACCTAAAACTAAAAAACTTACACAAGGCTATTGAAGATTCTTCTAAGGCACTGG 182

Qy       1055 AATTATTGATGCCACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGAC 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 AATTATTGATGCCACCTGTTACAGACAATGCTAATGCAAGAATGAAGGCACATGTACGAC 242

Qy       1115 GTGGAACAGCATTCTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGG 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GTGGAACAGCATTCTGTCAACTAGAATTGTATGTAGAAGGCCTACAGGATTATGAAGCGG 302

Qy       1175 CACTTAAGATTGATCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATG 1234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CACTTAAGATTGATCCATCCAACAAAATTGTACAAATTGATGCTGAGAAGATTCGGAATG 362

Qy       1235 TAATTCAAGGAACA 1248
              ||||||||||||||
Db        363 TAATTCAAGGAACA 376


Search completed: July  8, 2005, 12:03:37
Job time : 9626 secs
```